Using Nucleon Multiplicities to Analyze Anti-Neutrino Interactions with Nuclei

A THESIS SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Miranda Elkins

IN PARTIAL FULFILLEMTN OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE

Advisor: Dr. Richard Gran

July, 2017

Acknowledgements

This work was supported by NSF grant 1607381

Abstract

The most commonly used, simple interaction models have not accurately described the nuclear effects on either neutrino-nucleus or anti-neutrino-nucleus interactions. Comparison of data collected by the MINERvA experiment and these models shows a discrepancy in the reconstructed hadronic energy distribution at momentum transfers below 0.8 GeV. Two nuclear model effects that were previously not modeled are possible culprits of this discrepancy. The first is known as random-phase-approximation and the second is the addition of a meson exchange current process, also known as two-particle two-hole due to its result in two particles leaving the nucleus with two holes left in their place. For the first time a neutron counting software algorithm has been created and used to compare the multiplicity and spatial distributions of neutrons between the simulation and data. There is localized sensitivity to the RPA and 2p2h effects and both help the simulation better describe the data. Additional systematic or model effects are present which cause the simulation to overproduce neutrons, and potential causes are discussed.

Contents

List of Figures

List of Tables

Acronyms

QE Quasi-elastic

RPA Random Phase Approximation

MEC Meson Exchange Current

2p2h Two-particle Two-hole

MC Monte Carlo simulation

UMD University of Minnesota - Duluth

GENIE Neutrino interaction generator code

Geant4 Standard code that propagates particles through a detector

Low q3 Three momentum transfer $0 < q3 < 0.4$GeV/c

Mid q3 Three momentum transfer $0.4 < q3 < 0.8$ GeV/c

Chapter 1

Introduction

What neutrinos and anti-neutrinos do and how they interact with things around them has been essential input to studies of neutrino oscillations and the neutrino backgrounds to proton decay. Models have been developed to simulate what these two particles do; however, the models are not perfect and do not match high precision data on nuclei. To obtain a model that describes data, much still has to be done. In this thesis, two nuclear effects are extensively studied, and the models are compared to several observables. One observable in particular, the multiplicity of neutrons, is developed for the first time in MINERvA (probably a first for any GeV energy neutrino experiment).

The samples being studied to look for improved agreement between simulation and data are made up of anti-neutrino reactions on the protons in hydrogen and the protons and neutrons bound in carbon nuclei. The model currently being used is comprised of the basic Fermi gas assumptions where the proton is in motion in carbon and a single, average energy is paid for its removal from the nuclear (strong-force) potential.

The most basic reaction observed includes an anti-neutrino reaction with a proton which creates a neutron and an anti-muon. This reaction is referred to as quasi-elastic (QE) and its reaction equation is shown in Equation 1.1. The default QE simulation does not include a screening effect known as the random phase approximation (RPA). A second effect is due to a reaction that occurs at the moment two nuclei are exchanging a meson, the process that binds them together. This process is known as meson exchange current (MEC) interaction, or more generally a two-particle two-hole (2p2h) interaction. These reactions can be represented as shown in Equations 1.2 and 1.3.

$$\bar{\nu_\mu} + p \rightarrow \mu^+ + n \tag{1.1}$$

$$\bar{\nu_\mu} + p + n \rightarrow \mu^+ + n + n \tag{1.2}$$

$$\bar{\nu_\mu} + p + p \rightarrow \mu^+ + p + n \tag{1.3}$$

1.1 Three-Momentum, Energy, and Invariant Mass

The analysis focuses on the regions of low to medium three-momentum transfer in interactions between low-energy anti-neutrinos with carbon nuclei. Here low anti-neutrino energy is defined to be 2 GeV - 6 GeV; MINERvA has two higher energy anti-neutrino samples for potential followup studies. The low three-momentum transfer is defined as 0-0.4 GeV/c. Finally the medium three-momentum transfer is defined as 0.4-0.8 GeV/c. For brevity the low and medium three momentum transfer will be referred to as low q_3 and mid q_3 respectively.

Three-momentum transfer is a commonly calculated quantity in particle physics as is the energy transfer and the invariant mass of the system.

The quantity known as four-momentum is a four-vector in space-time that relates energy and three-momentum. The mathematical description can be seen in Equation 1.4.

$$P = \frac{E}{c}, P_x, P_y, P_z = \frac{E}{c}, \vec{p} \tag{1.4}$$

This is a quantity that can be calculated for any interaction but can vary between inertial frames. However, the square of this quantity is invariant, as is the square of most four-vectors, between inertial frames and is typically known as the invariant mass.

$$P^2 = (E/c)^2 - \vec{p}\cdot\vec{p} = (mc^2)^2 \tag{1.5}$$

The four momentum transferred (q) includes the energy and momentum carried by the virtual W boson in the simplest reactions such as the QE reaction discussed earlier. The Feynman diagram for the QE interaction can be seen in Figure 1.1 where the W boson is present. The four momentum tranferred in an interaction is defined by Equation 1.6.

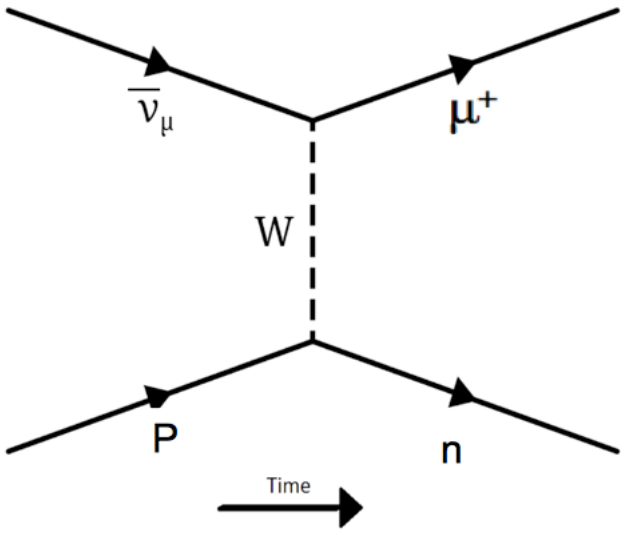

Figure 1.1: Example of a quasi-elastic anti-neutrino interaction with a nucleon.

$$q = P_{final} - P_{initial} \tag{1.6}$$

The square of the four momentum transfer is written in Equation 1.7. This quantity is also invariant. The neutrino, target proton, and target nucleus (the latter is effectively "lab" reference frame) can be thought to each have their own reference frame that is different from one another. The same can be said for the outgoing anti-muon and recoil neutron. They will disagree on how much momentum and energy was transferred but they will agree on one value of q^2.

$$(P_{final} - P_{initial})^2 = (E_{final} - E_{initial})^2 - (\vec{p}_{final} - \vec{p}_{intial})^2 = q_0^2 - \vec{q_3}^2 \tag{1.7}$$

Here the energy transfer is given as the variable q_0 and the three-momentum transfer is given as the variable $\vec{q_3}$ as before. The notation for this will be simplified to q_3 meaning the magnitude of the three-momentum vector and not solely the third component of the four-momentum vector.

Calculating three-momentum transfer is typical of any type of particle interaction. The equation used to define this in the anti-neutrino interaction inside the carbon nucleus is elementary.

$$q_3 = \vec{q_3} = \vec{p_\nu} - \vec{p_\mu} \tag{1.8}$$

Here the three-momentum transfer is determined by the vector difference between the magnitude of the three-momentum of the anti-neutrino and the anti-muon. This is problematic because the anti-neutrino is invisible to the detector and thus its three-momentum must be obtained by other means.

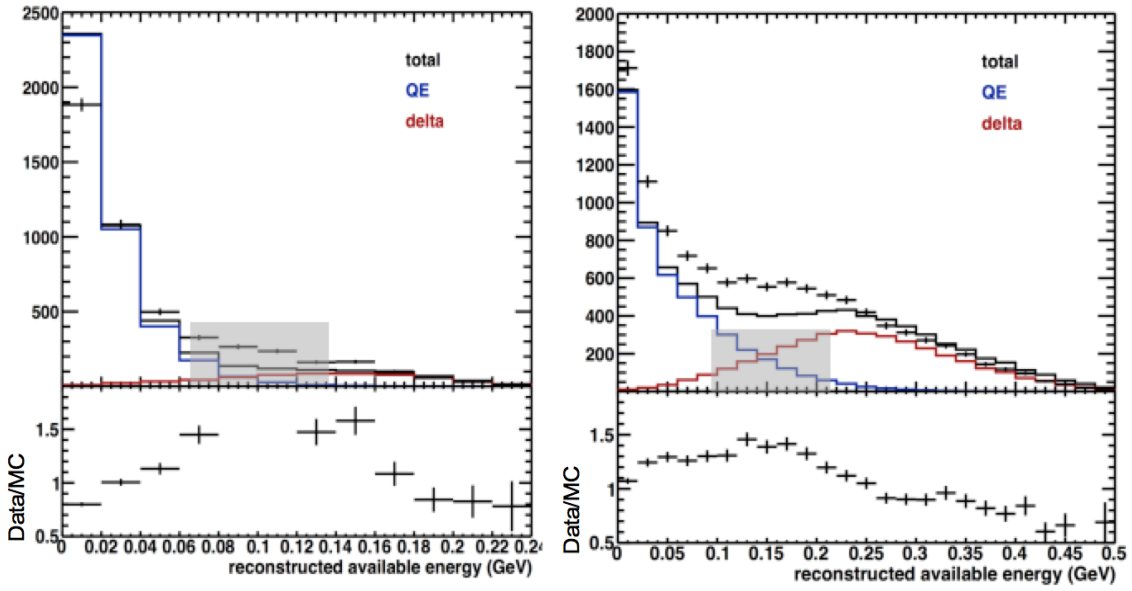

Figure 1.2: Reconstructed available energy without the RPA re-weight or the 2p2h events added. Low q3 is shown on the left and mid q3 is show on the right. Separated into three hadronic energy regions: QE before shaded region, Dip inside shaded region, and Delta after shaded region.

The plot shown in Figure 1.2 depicts the reconstructed hadronic energy in the tracker portion of the MINERvA detector for real data along with what is expected from simulation. The differences between the model and the data can be best identified by this plot and therefor it will be referenced throughout the paper. In the plot the two interaction types being modeled (QE and Delta) can be plotted as subcomponents of the simulated events for simulated events. The 2p2h processes are expected to be most apparent between the QE reactions and delta resonance (historically called the "dip region") due to the anticipated hadronic energy associated with each.

1.2 Quasi-Elastic

Figure 1.2 is the elastic reaction, whose peak is at the lowest reconstructed hadronic energy. The interaction equation is repeated as Equation 1.9. Quasi-elastic interactions nearly conserve kinetic energy similar to how a completely-elastic interaction, like billiard ball collisions, would completely conserve kinetic energy. This near conservation allows quasi-elastic interaction kinematics to be calculated in the same way as a completely-elastic interaction would be.

$$\bar{\nu_\mu} + p \rightarrow \mu^+ + n \tag{1.9}$$

The Feynman diagram of this interaction is shown in Figure 1.3, which has also been repeated. Here an anti-muon-neutrino interacts with a proton via the exchange of a virtual W^- boson and produces a neutron and a μ^+. The physics of these particle interactions demands the momentum, energy, charge, and the lepton and baryon numbers all must be conserved.

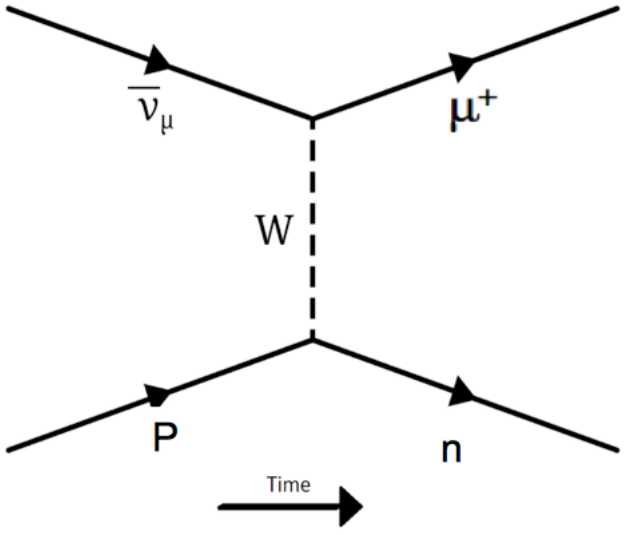

Figure 1.3: Example of a quasi-elastic anti-neutrino interaction with a nucleon.

Calculating the full kinematics is done using a technique not specific to QE events. We start with the energy of the neutron produced in the interaction along

with the muon's energy and angle. From this we construct the value of the anti-neutrino's energy and then q^2 where the equations for this and $q3$ are repeated in Equations 1.10 and 1.11.

$$q^2 = q_0^2 - \vec{q_3}^2 \tag{1.10}$$

$$q_3 = \vec{q_3} = \vec{p_{\bar{\nu}}} - \vec{p_\mu} \tag{1.11}$$

The problem with this is neutrons are neutral and have a long interaction length in comparison to the detector. Thus they typically do not interact often within the MINERvA detector, making reconstructing their energy difficult which leads to a bias, too low reconstructed hadronic energy and q3.

In the neutrino case, instead of neutrons being the products in our QE interaction one would find protons instead. The energies, and sometimes the angles, of the protons produced are measured along with the muon's angle and energy. Then through the relativistic expression for energy, momentum and mass the other can be calculated. Because they are charged particles, protons and their energies are more easily measured, which makes the neutrino version of the quasi-elastic process simpler to analyze.

Current and next generation oscillation experiments are searching for CP violation by looking for distortions of the neutrino energy spectrum that are different for neutrino and anti-neutrinos. But here intrinsically the reconstructed energy spectrum is different just because of neutrons in the final state. For these oscillation experiments to succeed, we need dedicated experimental data and models that describe the production of these protons and neutrons accurately.

1.3 Delta Production

The next type of interaction that is part of the model is known predicted by the model is known as the "Delta Resonance". The reaction equations can be seen in Equation 1.12 and Equation 1.13.

$$\bar{\nu_\mu} + p \rightarrow \mu^+ + \Delta^0 \tag{1.12}$$

$$\bar{\nu_\mu} + n \rightarrow \mu^+ + \Delta^- \tag{1.13}$$

The delta can be the excited state of a neutron or a proton. A better description is a delta particle is comprised of some combination of three up or down quarks, where its charge is determined by these quarks. An up quark has a charge of $+2/3$ that of an electron and a down quark has a charge of $-1/3$ that of an electron. The uud (up,up,down) and udd (up,down,down) quark states for the delta particle have the same charge as for the proton and neutron respectively. The one in Equation 1.13 is negatively charged and is made of a ddd (down, down, down) combination. There is a Δ^{++} made of uuu (up,up,up) also, which can not be produced by anti-neutrino reactions. The difference is the delta particle is an excited state with 0.3

GeV more energy than either the proton or the neutron. The Feynman diagram for one of the less common delta interactions we see involving anti-neutrinos can be seen in Figure 1.4 [11].

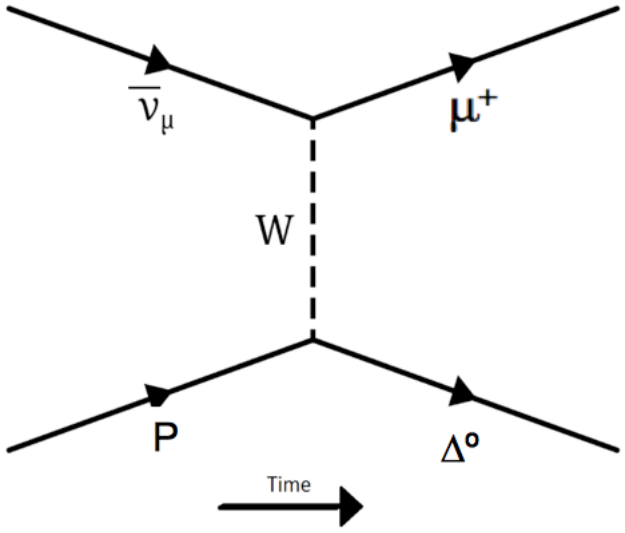

Figure 1.4: Example of a delta produced from an anti-neutrino interaction with a nucleon.

The deltas are a temporary excited state and they will quickly decay, sometimes before exiting the nucleus, into a nucleon and a pion. For the decay of the neutral delta it would be typical to see a neutron and a neutral pion (pi zero) or a proton (less likely) and a negatively charged pion (pi minus). The decay of the delta minus would produce a neutron and a pi minus. These interactions require a neutron for anti-neutrino reactions, and only occur in carbon nuclei in the detector. It is also possible for the Δ^0 to decay into a neutron and a photon, but this is rare.

If the process happened in the carbon nucleus, the decay could happen so fast the Delta hadn't time to leave the nucleus. The delta and/or its products can re-interact with another nucleon before leaving the nucleus, known as intra-nuclear re-scattering. In this case the final states would be modified from what is described above. A particularly interesting case is when the pion is "absorbed" by the nucleus here meaning re-scattered by a nucleon within. That nucleon now leaves the nucleus, instead of the pion, along with the other nucleon that was paired with the pion as the original products of the interaction. This allows for a two nucleon final state coming from the delta production which looks very similar, in terms of final state and available hadronic energy, to the 2p2h interaction discussed in section 1.5.

1.4 Random Phase Approximation

Random phase approximation [7], [1] is a method to compute a multi-nucleon effect caused by the existence of nucleons other than the one participating in the anti-neutrino-nucleon interaction. In RPA the nuclear potentials are modified which

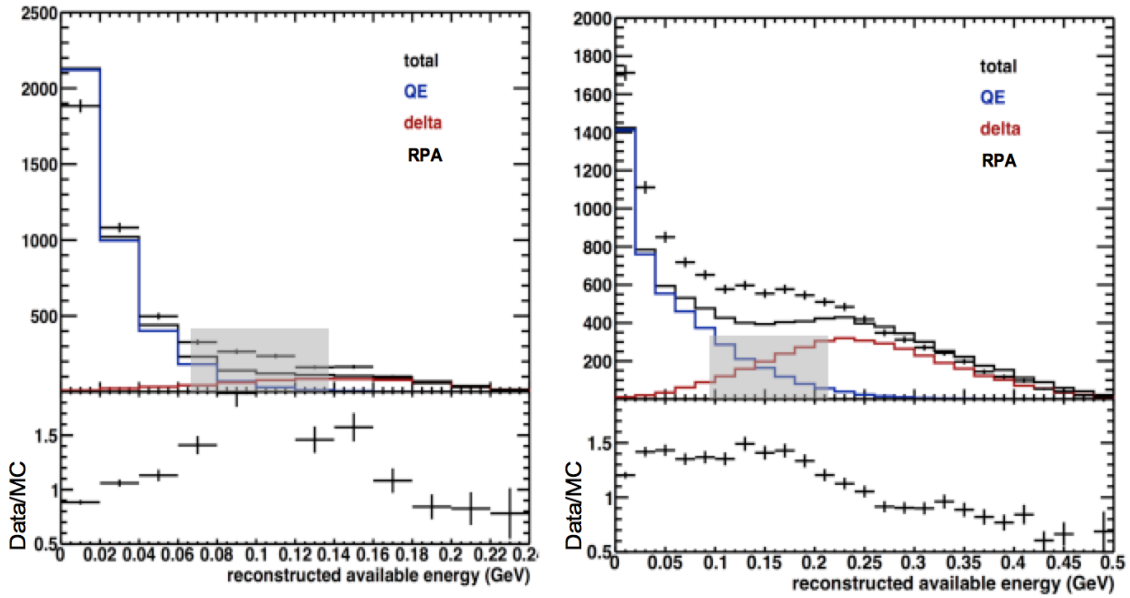

Figure 1.5: Reconstructed available energy with only the RPA re-weight added for the low q3 (left) and mid q3 (right).

effectively changes the energy levels of the nucleus and also causes a decrease in the probability of an interaction with one of these nucleons. Often it is called a screening effect, and has analogies in classic polarization screening. This suppression causes a visual difference between the plots shown in Figures 1.2 and 1.5 The addition of this effect causes a suppression of the QE resonance, most noticeable in the first bin, for both low and mid q_3 but is more pronounced for the low q_3.

The neutrino version of this analysis [12], [9], [7] was the first to isolate the need for a RPA suppression in neutrino scattering. Likewise, the study in this thesis aims to be the first to do so for anti-neutrino scattering.

1.5 Two-Particle Two-Hole Interactions

The final reaction of interest, nicknamed 2p2h, won't happen for hydrogen, and is only possible in nuclei. The two reaction type equations are shown in Equation 1.2 and Equation 1.3. The Feynman diagrams of these two interactions can be seen in Figure 1.6.

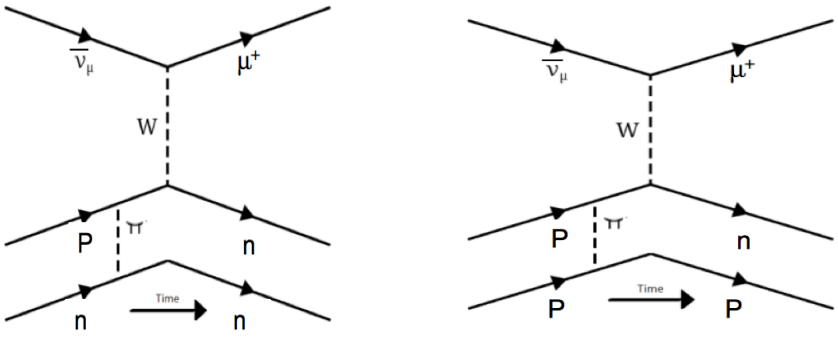

Figure 1.6: Example of a meson-exchange-current interaction with a neutron-neutron (left) and a neutron-proton (right) final state.

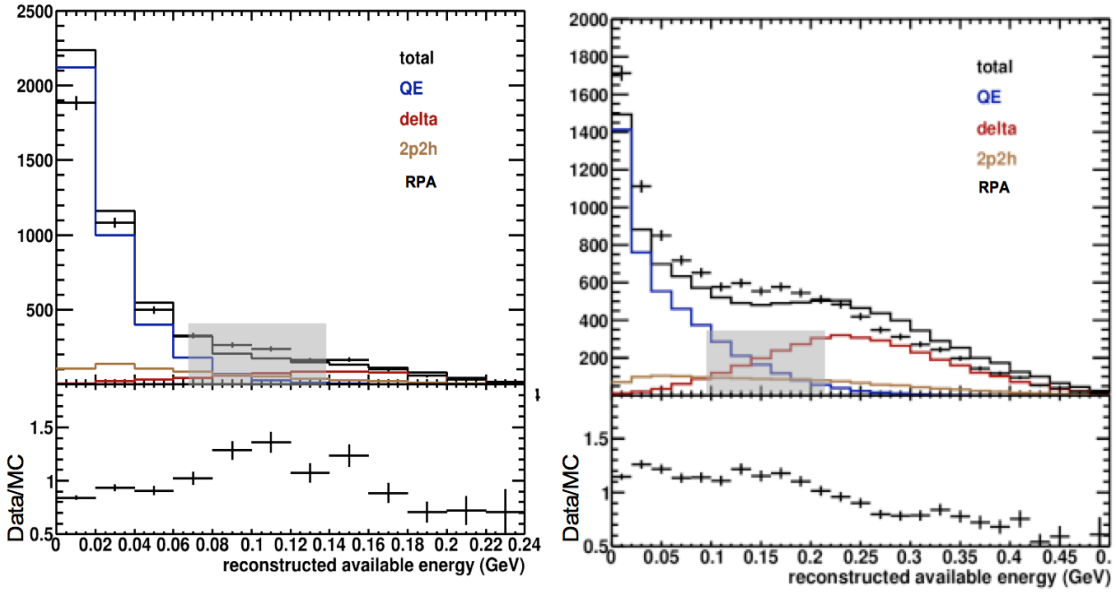

Figure 1.7: Reconstructed available energy with both the RPA re-weight and the 2p2h added.

The Valencia model [2], [3] for 2p2h reactions predicts a range of energy and momentum transfer, including in between the delta and the QE process. In Figure 1.2 one can see that in this region, nicknamed the "dip region", the model poorly describes data which indicates we are not modeling correctly and this may be the most important missing model element.

The 2p2h process explicitly has two nucleons knocked out of the nucleus. The two nucleons are necessarily either a proton-neutron pair (from a proton-proton initial state in carbon) or a neutron-neutron pair (from a proton-neutron initial state). Ordinarily, the QE and Delta reactions will lead to a single nucleon, most often a neutron. Except during the instances where the pion from the delta interaction is absorbed, or the neutron from a QE knocks out another nucleon, allowing for two nucleons to leave the nucleus. Detecting and counting nucleons may reveal the

otherwise unique signature of the 2p2h process, and is the primary motivation and goal of this thesis.

If we add in the newest model predictions for the 2p2h interactions we can see the new events appear between the QE interactions and the delta resonance and the model, to the eye, now more accurately fits the data. However, there are apparent discrepancies that still exist.

1.6 More Motivation

The development of new neutrino oscillation experiments and new detectors will rely on our ability to understand the different parts of the neutrino and anti-neutrino interactions especially their products as the incident neutrinos and anti-neutrinos are invisible to us. The unbiased reconstruction of the neutrino's or anti-neutrino's energy and momentum requires a precise understanding of neutron production, especially for the anti-neutrino, but the uncertainty in neutrons being simulated is quite large.

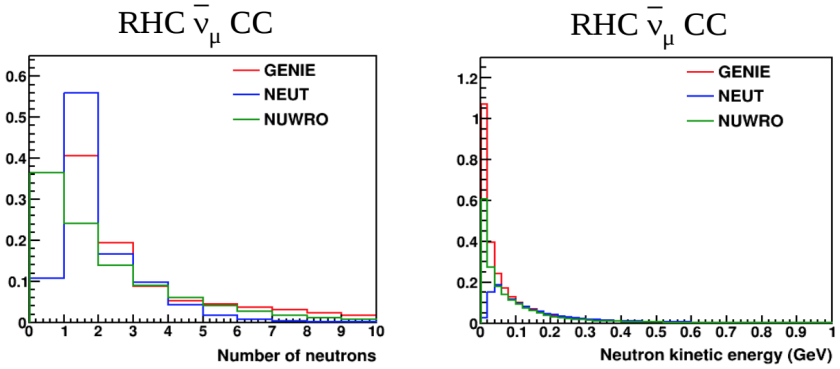

Figure 1.8: Left: The number of neutrons produced during anti-neutrino interactions simulated by the three different Monte Carlo (MC) event generators. GENIE is featured in this paper. Right: The number of neutrons for a given neutron kinetic energy produced by each event generator. GENIE produces the most low energy neutrons and the most neutrons on average.

There are various neutrino event generators used by the high-energy physics community, three of which include GENIE, NUWRO, and NEUT. The differences between these three event generators are great when it comes to how they produce neutrons in both neutrino and anti-neutrino simulated events. Figure 1.8 shows a short study completed by several collaborators from the DUNE collaboration [8]. Here the number of total neutrons produced by each event generator and the kinetic energy for each of the neutrons produced are plotted. GENIE tends to produce the most neutrons as well as the most low energy neutrons. This analysis uses GENIE as the event generator and aims at identifying whether or not these large numbers of low energy and overall neutrons match what is produced in data.

1.7 Analysis Overview

Given Figure 1.5 and Figure 1.7 it is easy to assume the additions of the RPA effect and the 2p2h interactions help our models achieve a better fit to the MINERvA data in most regions of hadronic energy. It is obvious from the hadronic energy plots these two effects being added are not the entire story.There are other observables we can use for comparisons between data and simulation. Identifying the 2p2h interactions associated with the anti-neutrino case will proceed with an attempt to reconstruct neutrons accurately. A major component of this work is the development of a new software algorithm created to identify neutrons in the MINERvA data. The tool utilizes C++ code along with information that can be extracted from the MINERvA data and simulation for each reconstructed anti-neutrino interaction event.

The algorithm focuses on identifying neutrons based on energy and location of the interaction within the MINERvA detector. The low and medium three-momentum transfers, 0.0 - 0.4 GeV and 0.4 - 0.8 GeV respectively, which allows for the elimination of more active interaction types that would make this type of study more difficult.

With the finished algorithm the neutron candidates, possible neutrons identified by the algorithm, are then counted in each anti-neutrino interaction event with specific selections on momentum transfer and hadronic energy. The multiplicity of neutrons will be compared to scenarios with different amounts of the three reaction types QE, 2p2h, and Delta. This is done for both simulation and data as to determine the differences between the two in hopes of identifying the cause of the discrepancies seen in the reconstructed hadronic energy plots, and what effects can be used to create a better agreement.

There is a previous algorithm written for the neutrino case used to identify proton candidates in an event. This is yet another observable we can use to fully understand the MINERvA anti-neutrino data and the differences between it and the simulation. This algorithm, written by John Demgen, with help from Alec Lovlein and Jake Leistico, is reused here for the anti-neutrino sample to create proton multiplicity distributions.

Chapter 2

MINERvA and MINERvA Analysis Tools

The MINERvA detector at Fermilab is the provider for the anti-neutrino data being analyzed in this paper. This section contains the description of how the neutrinos and anti-neutrinos interacting in MINERvA are detected and how the detector is structured. To enhance the description of the detector the event display Arachne is described and used to visualize events.

For this analysis a neutron counting algorithm was written to help aid in the identification of nuclear effects in the anti-neutrino data taken by the MINERvA detector. It became clear during the algorithm's development that certain tools already created by previous MINERvA collaborators would be useful. The tools used are known as a muon fuzz eliminator, a tracked particle identifier, and a vertex blobber. A description of each of these tools is also in this chapter.

2.1 MINERvA Detector

The MINERvA detector is strategically placed in the NuMI beam line at Fermilab where neutrinos and anti-neutrinos with a spectrum from 2 to 20 GeV give the collaboration a wide range of neutrino and anti-neutrino interactions to examine. To create the NuMI beam at Fermilab, a beam of 120 GeV protons from the main injector strikes a graphite target producing pions and kaons. These charged particles are then focused by the magnetic field of two focusing elements (called "horns") which direct the pions and kaons into a decay pipe where they do decay as anticipated by the name. These focusing horns can be set in two polarities with a change of the direction of the 185 kA current. When the horns are focusing positively charged pions and kaons the resulting beam is primarily neutrinos, and when the horn is focusing negatively charged pions and kaons the beam is primarily anti-neutrinos [4]. These beams are not entirely neutrino or anti-neutrino as there is some contamination within both of the beams from particles of the opposite type.

The MINERvA experiment was designed to look at neutrino-nucleon interactions with various nuclei. Due to this the design of the detector required many key features which are depicted in Figure 2.1. This analysis uses only a small subset of those

features.

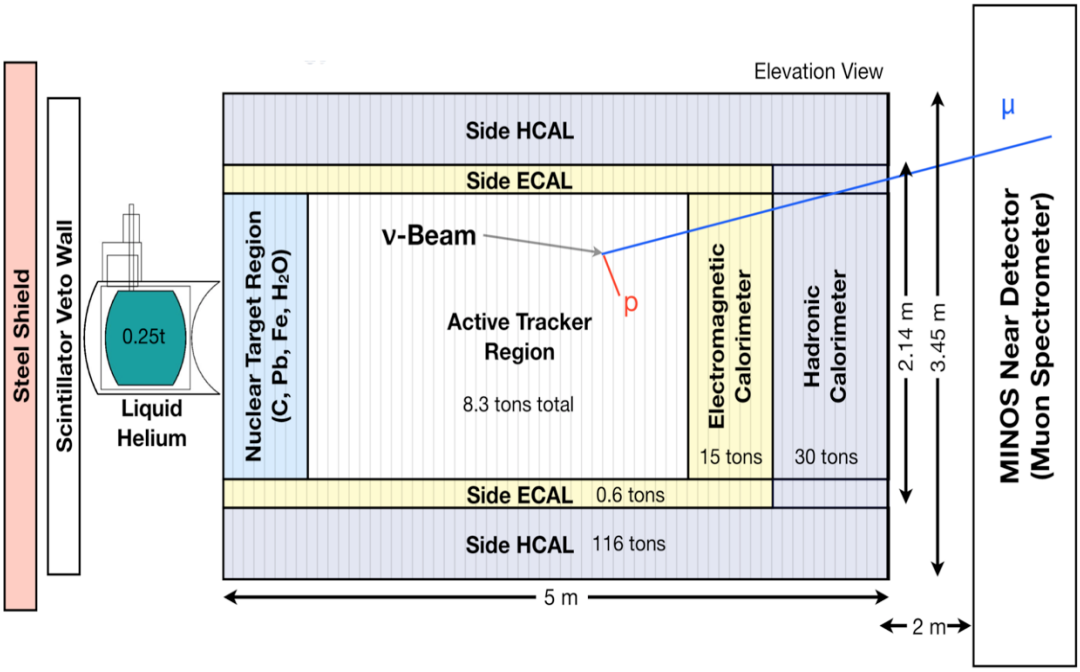

Figure 2.1: Schematic of the MINERvA detector [4]

The MINERvA detector is separated into over two hundred hexagonal shaped segmented scintillator planes which are stacked like hanging file-folders to form an inner detector with an outer side detector. The inner detector contains a nuclear target region, the tracker region, the electromagnetic calorimetry (ECAL) and hadronic calorimetry (HCAL). The nuclear target region includes a mixture of scintillator along with iron, lead, and water.

A veto wall made up of two alternating sets of a steel plate and a scintillator plane, which illuminates when excited by ionizing radiation, is placed the farthest upstream, the first piece of the apparatus encountered by the anti-neutrino beam. This veto wall allows for background particles including hadrons and muons coming from the rock wall which the beam is passing through to be tagged and characterized so they are not accidentally included with the products of reactions in the center. From here the beam goes through a cryostat filled with liquid helium before entering the main part of the detector. The scintillator planes each contain 127 scintillator strips with an optical fiber inside. These fibers allow the light caused by the ionizing radiation to travel towards PMTs where it is collected and then turned into quantities measurable by the MINERvA collaboration.

The anti-neutrino reactions for this study are selected to be from the tracker region. The tracker region is also where most of the neutron products of those reactions are found and makes the highest resolution measurement of the neutrino reaction products. The region is made up of planes of the scintillator described above. A group of two planes together are known as a "module" where the planes

in each module have two different orientations of the three possible. These three different orientations are known as X, U, and V where the X planes are vertical and the U and V orientations are rotated 60 degrees either clockwise or counter-clockwise respectively. Each module has one plane in the X view and the other in either the U or V view. The modules alternate between this UX or VX structure with the planes in the X view always more downstream, or further from the direction of the incoming beam. The tracker has 62 of these modules.

This module structure found in the tracker region is exactly the same for the ECAL as well. The difference here is that each plane of scintillator has a 0.2 cm layer of lead glued to it. The side ECAL also is composed of lead but here it varies in thickness. There are ten modules in the ECAL (twenty planes). The HCAL scintillator planes make up a different pattern then discussed for the other three regions of the detector. Here there are 20 different modules, each made up of a single plane of scintillator and a one inch think plane of steel. The scintillator planes have a repeating orientation pattern of XVXU which is similar to what is found in the tracker and ECAL except now with the steel planes in between each.

2.2 Arachne

The neutron counting pattern recognition algorithm was developed (and is easiest to describe) with a visual event display known as Arachne. The event display is a web based tool created by Nathaniel Tagg (professor at Otterbein College) and used by the MINERvA collaboration by the MINERvA collaboration that allows the user to visualize what a neutrino or an anti-neutrino event looks like inside the detector [14]. Visualizations can be made for both data and simulated events. The latter can include the extra information about what "truth" situation was simulated. This information includes a visual of the detector split up into the modules in each of the three views, particle parentage trees, time lines, energy size of particles and deposits within the detector, plus much more. Examples of the most used portions of the web display are shown through out this section.

Working from the top of the Arachne web page down the first GUI below the Arachne title, named Data Source, allows the user to choose what event file they would like to look at. All of the MINERvA data and simulation events are split up by a run number, then by a subrun number, and even further by a gate number. An example of a run number is 50201, which is part of the anti-neutrino data used for this analysis. An example of this is seen in Figure 2.2.

One pulse of neutrinos produces several neutrino interactions in MINERvA. As long as their separation in time is more than the 150 ns, the duration of time where information is collected for an event, it is easy to isolate a single one. Figure 2.3 is an example of what the next portion of the event display looks like. The box titled Time Histogram shows how much activity there is in the detector as a function of time for each event. The histogram is split up into "slices" of time and the small box located just above allows the user to skip ahead by one slice at a time. As this occurs the rest of the boxes update their information to correspond to the activity which occurred during the currently viewed time slice.

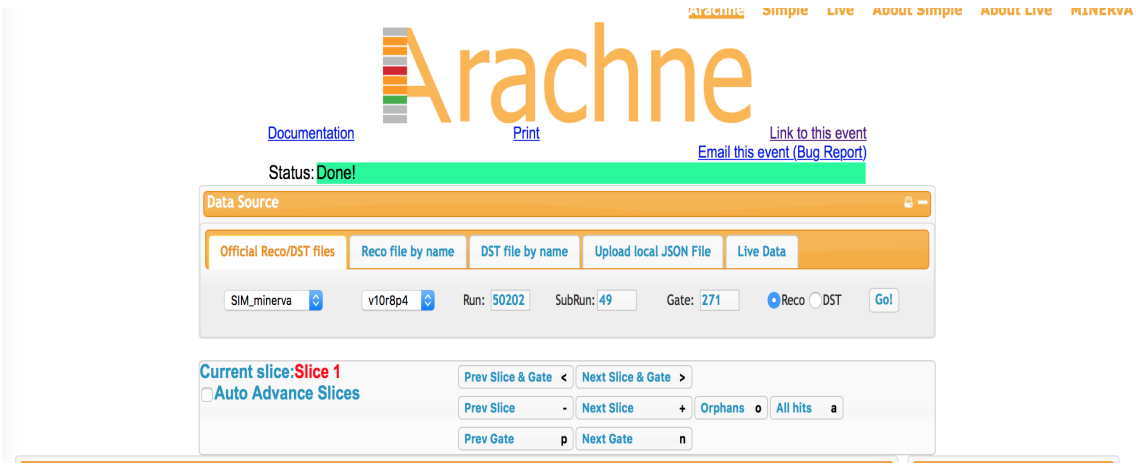

Figure 2.2: Event from MC run-50201, subrun-49, and gate-271. Arachne display data source information. Link to the event: `http://minerva05.fnal.gov/Arachne/arachne.html?det=SIM_minerva&recoVer=v10r8p4&run=50202&subrun=49&gate=271&slice=1`

Neutrons from the anti-neutrino interactions used for this analysis tend to be low in energy, less then 100 MeV. These neutrons re-interact and deposit small amounts of energy, usually much less then 20 MeV per strip, whenever they do re-interact. This makes the box to the right of Time Histogram, named PH Histogram, useful because it allows the user to adjust the the color of the energy scale of the energy deposits so they are easier to view in the hit map boxes. For visual neutron hunting, the scale used was normally between .5 MeV and 20 MeV so that smaller changes in energy deposits were more apparent.

The bulk of the information utilized for neutron hunting came from the hit map boxes used to actually view the energy deposits in the event. There are two hit map boxes, one titled Hit Maps that splits up the 3 views of the inner detector and 6 outer parts of the detector plus one titled Big Hitmap which shows a larger display of the three views of the inner detector. The former gives the overall picture, the latter is good for looking at fine details. This first can be seen in Figure 2.3 while the second can be viewed in Figure 2.4.

If you imagine these are like digital photographs, these two hit maps show each individual pixel as the triangles they really are.Each represents one scintillator strip in the three views. Each strip has a color corresponding to the energy deposited, where the color scale is shown in the PH HIstogram box. This example is of an anti-neutrino event interaction with carbon. Leaving the nucleus in the interaction are a negatively charged pion, a proton, and a neutron. The black lines indicate where each particle traveled within the detector.

The goal is to design an algorithm to identify neutrons in the data. The key is to understand how neutrons deposit energy in the simulated detector. In a way we are trusting the simulation of neutrons in order to infer how many come from several categories of real neutrino interactions. Using truth information for simulated events it was possible to do just this. The Slice Info and Monte Carlo boxes to the right of the hit maps contain much of the information about the interaction. The particles

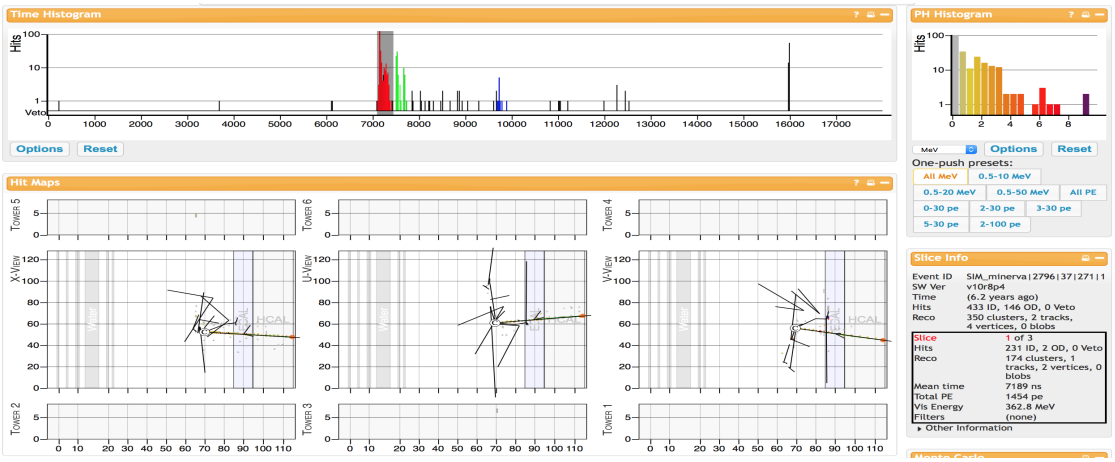

Figure 2.3: Event from MC run-50201, subrun-49, and gate-271. Arachne hit maps and energy distribution information

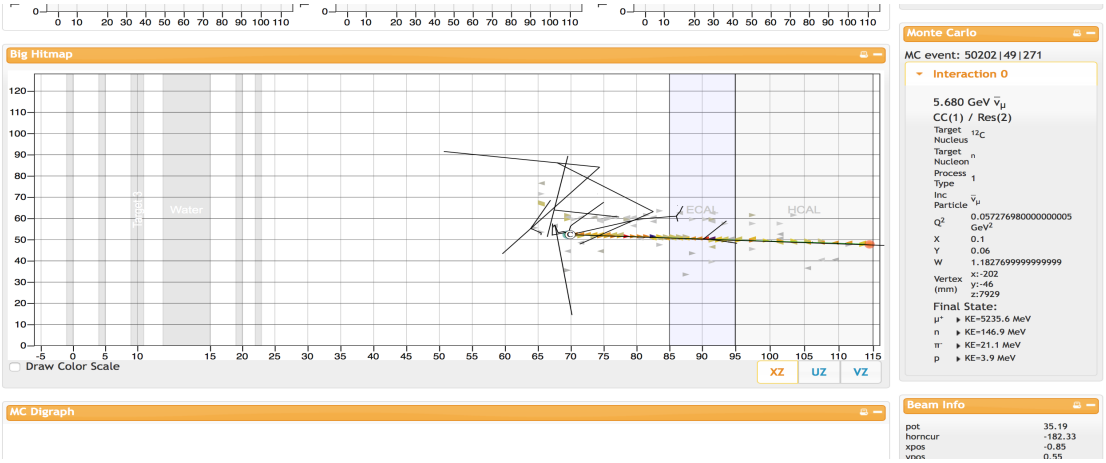

Figure 2.4: Event from MC run-50201, subrun-49, and gate-271. Arachne big hit map and MC information

came from the GENIE interaction simulation and given to the Geant simulation, along with their energies, where the vertex of the interaction was, the length of time elapsed, plus much more is included. Of course the Monte Carlo box is empty if the event being looked at is from the MINERvA data set and not simulation.

An important distinction for the neutron analysis is what deposits are from particles re-interacting that came from the nucleus (GENIE particle) of the anti-neutrino interaction and which deposits are caused by re-interaction particles also re-interacting with other nuclei in the detector (Geant4 particle).The MC Digraph portion of Arachne gives us this information and, like the Monte Carlo box, would also be empty for data. Here the particle parentage tree, shown in Figure 2.5, says what each particle from the original interaction is along with the products of its own interactions and so on for several generations. If a user clicks to highlight a certain particle the truth line of the clicked particle will appear solely on the hit maps. These lines show where the particle, and its parents, moved throughout the

detector. Any clusters of energy beneath a truth line would indicate some type of interaction occurring. This information was used to validate many aspects of the neutron identification algorithm.

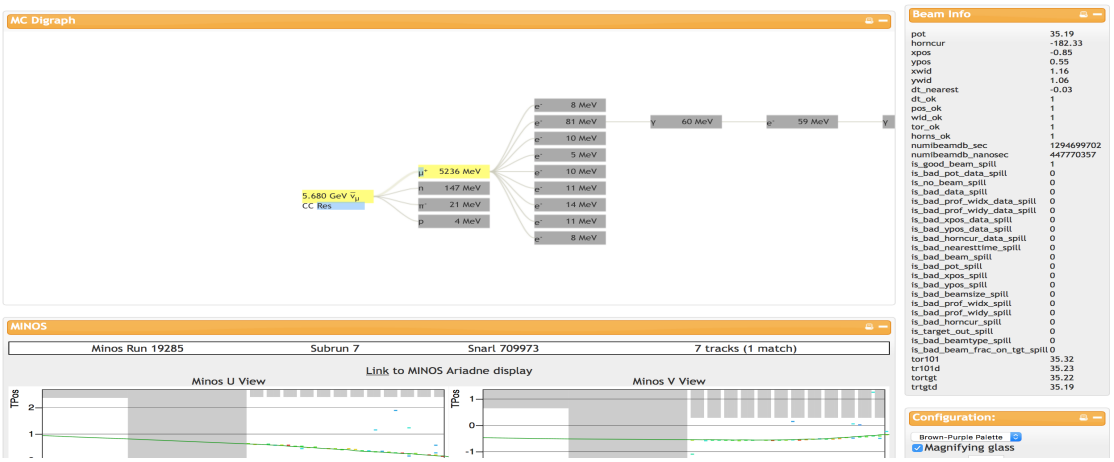

Figure 2.5: Event from MC run-50201, subrun-49, and gate-271. Arachne particle heritage information

The Arachne event display has quite a bit more information to offer then was utilized and relevant for this analysis. All of the information about the beam creating the neutrinos and anti-neutrinos at the time of the currently viewed interaction is provided. The current in the horn, the beam position in 3 dimensions, the number of protons on target plus much more is included.

Another part for users to view is the MINOS box. MINOS is another detector at Fermilab that is positioned next to the MINERvA detector. The muons from the neutrino and anti neutrino interactions are sometimes also detected through MINOS and are called MINOS matched. This is also an interactive 3D display that allows a person to rotate the MINERvA detector in any direction with all of the activity from the neutrino interaction being displayed all at once. There are tabs within that zoom in and out, switch views, and move the detector display in 4 directions without the user having to do so on their own. The hits within MINOS can also be toggle on and off to be displayed here.

2.3 Muon Fuzz Eliminator

The total energy of the interaction products excluding the muon allows us to estimate the neutrino energy, it is the estimate of the energy transfer, and is needed to estimate the three-momentum transfer. As an example of energy coming from the muon that could accidentally be counted as hadronic energy, it is common for a wide variety of photons to be emitted from the muon through a process called Bremsstrahlung. Also the muon could strike an atomic electron and give it an energetic kick, called a “knock-on electron”. Both these can produce activity that is separated from but usually close to the muon track. One of these occurs in the event displayed in the previous section; easily visible in the middle of the ECAL.

The energy they deposit is obviously from the muon, meaning we do not want to include it while reconstructing the available hadronic energy, and this known as "muon fuzz". The muon fuzz is not perfectly simulated and work is currently being done to characterize this uncertainty.

To ensure the energy summed from the products does not include anything from the muon or anti-muon re-interactions a tool was developed to help ignore the energy deposited in the detector that came directly from this muon or anti-muon. The muon fuzz tool creator is Lu Ren, who recently defended and published her thesis with MINERvA data [13]. The tool identifies where the (anti-)neutrino first interacts in the detector creating an (anti-)muon and whatever other products are created with it. This is known as the vertex. The tool then tracks the energy deposits created by the (anti-)muon in the detector and draws something similar to a box around it that is wide enough to enclose most of the clusters caused by products of (anti-)muon re-interactions. Anything within this box is then determined to be either the muon or "muon fuzz" and ignored for later with energy reconstructions.

2.4 Tracked Particle Eliminator

In the MINERvA detector high-energy charged particles such as protons and charged pions tend to leave deposits of energy that appear to look like a track from the primary interaction point. This is something neutrons never do so we should certainly never form neutron candidates from them. The tracked particle identifier is useful for doing just this. The tracks caused by charged particles begin right at the vertex and usually extend in a direction different from the track one would see from the muon or anti-muon. These tracks differ from that of the muon because they are typically not nearly as long or cross through as many strips in the detector. The tool is capable of identifying energy clusters close enough to one other that extend significantly far into the detector, yet not being from the muon or anti-muon, from the vertex and tag these to be "tracked".

2.5 Vertex Blobber

The final and most important tool being utilized in the algorithm is the vertex blobber. The tool blobs clusters together near the vertex of a neutrino or anti-neutrino interaction. The term blob means to add clusters of energy together found in the strips of scintillator of the MINERvA detector if they are near enough to one another and fit some other spatial criteria. These criteria can be set by the user based on the what it is they need the tool to do. For this analysis, the goal for this tool is to eliminate all activity that is simply-connected to the vertex of the anti-neutrino interaction. Because neutrons have a relatively long interaction mean free path compared to the size of the detector, activity near the vertex is far more likely to be from charged particles like protons and pions.

In the tool there are several quantities which need to be defined and questions that need to be answered such as: three radii are required to be specified in mil-

limeters, a list of the clusters that can potentially create a blob is needed, and it has to be determined if the clusters should be filtered or not. Filtering the clusters causes each cluster to be ignored if it is heavy ionizing, low activity, or cross talk. Every cluster has a flag that identifies whether it is used or not. Being used means the cluster sits on the muon track or is already being used by the vertex blobber and therefor it would be ignored.

There are two modes in which the vertex blobber can be used; single blob or filament blob and the modes are determined by the three radii that are set by the user. The first radius is the maximum search distance the user wants to look for clusters from the vertex that can be potentially added to a vertex blob. This distance could be made large enough to encompass the entire detector. The second radius is the maximum starting distance a cluster must be within the vertex for it to be able to start a vertex blob. The final radius is the maximum gap distance between any cluster already a part of a vertex blob and a new cluster that can be added. If the user sets the first radius to be where he or she wants to search and create blobs and the second two radii larger than or equal in distance to the first, a single large blob will be created with all clusters within that first set distance, essentially making a simple sphere-like box around the vertex. If the user sets the first radius to some moderate size but makes the starting distance much shorter and the gap distance small a filament blob will be created. A filament blob will follow simply-connected energy deposits out to an arbitrarily large distance.

The single blob mode is the default mode for all other MINERvA analyses as of this writing, but it is not adequate for the neutron counting algorithm. The filament blob mode was developed by Jyotsna Osta and Gabe Perdue, and revalidated for use here. In my use-case, I want to the algorithm wants to eliminate clusters near the vertex and clusters on a track that extend away from the vertex but aren't caught by the tracked particle identifier. Due to these tracks extending quite a distance sometimes, if the single blob mode was used with such a large radius it is possible to start adding actual neutron clusters into the blob which would then be ignored and never counted as a candidate. The opposite is also true where if the radius is set to be too small and a tracked particle interacts and skips a plane, causing the last energy deposit to not be tracked, we get extra neutron candidates.

There is also a catch with the filament blob mode which needs to be addressed. The filament blob mode only creates a single blob when the tool is turned on. The tool will loop through the list of clusters provided and will start a blob with the closest cluster to the vertex within the user's set starting distance. The tool will then add any additional clusters that are within the gap tolerance between clusters that can be added to a blob. The tool will then mark all of these clusters as used and then stop. It is possible for the there to exist a cluster which is not near enough to the just created blob to be added to it but close enough to the vertex to start a blob of its own. Due to this the tool must be run iteratively to ensure all of the activity near the vertex is blobbed. All of the clusters added to the vertex blobs from all iterations are then flagged so the algorithm can deem these as non-neutron candidates.

Examples of both one and two filament blobs being created are shown in Figure

2.6 and Figure 2.7. The first Arachne display shows two different filament blobs, one above the interaction vertex and one below the interaction vertex. It is easy to see the top cluster was too far away from the original blob created and had the tool not been run iteratively it would have not been tagged to be in a vertex blob. This means it eventually would have been counted as a neutron candidate which is not what is needed for the algorithm to perform optimally. Had the blobber been run in single blob mode with a set blobbing distance large enough all of the circled clusters would be considered the same vertex blob.

It is common for a tracked particle to skip a plane and produce a cluster off of the track seen in the detector. The second display gives an example of what this would look like. The filament blob is ideal to catch these types of clusters, and the neutron finding algorithm won't count them as a neutron candidate. Had there been a cluster at this distance but not on a track the algorithm would want to count it as a neutron candidate, but clearly it is some other charged daughter particle from an interaction. So this also helps demonstrate why the single blob mode would start eliminating some of the neutron candidates we do want and shows why the filament mode is best for this situation.

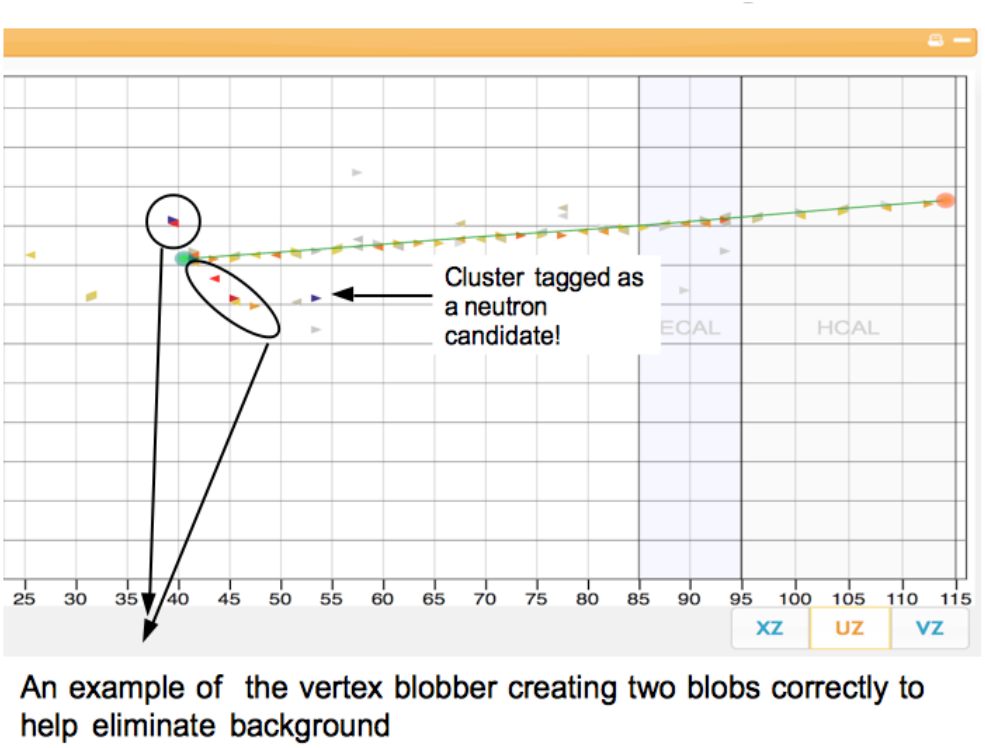

Figure 2.6: Event from MC run-50201, subrun-152, and gate-1063. The event includes a collection of clusters surrounding the vertex of the interaction. These clusters are circled where the vertex blobber created two different blobs which eliminated these clusters from creating neutron candidates.

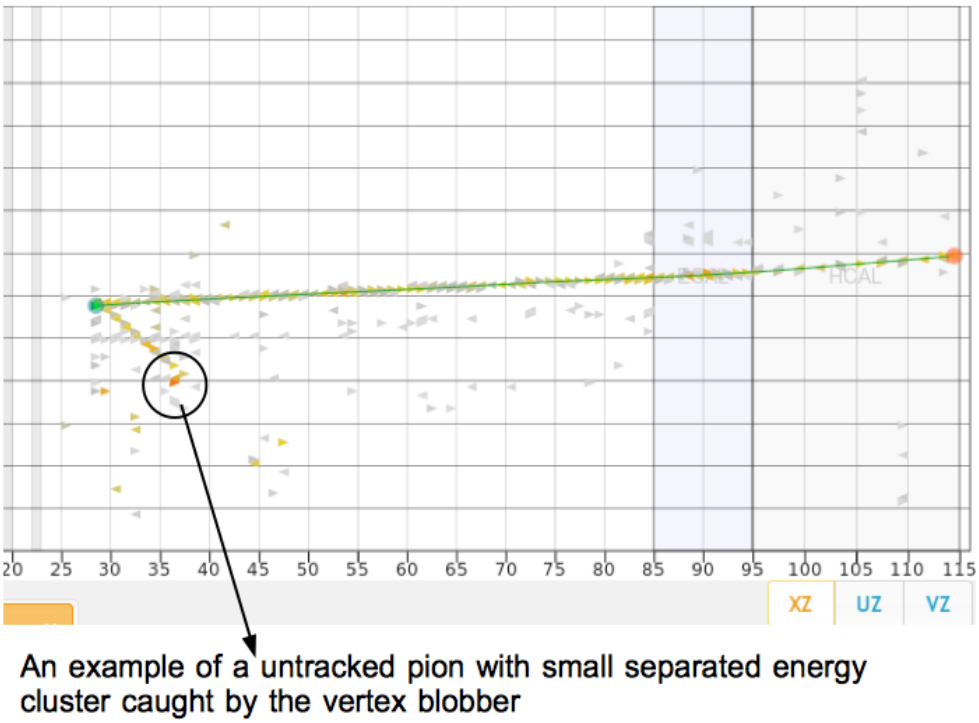

Figure 2.7: Event from MC run-50201, subrun-160, and gate-935. The event includes a collection of clusters surrounding the vertex of the interaction. These clusters are circled where the vertex blobber created one filament blob which eliminated these clusters from creating neutron candidates.

The specific parameters used and validated for this analysis are as follows. The maximum search distance is set to be 2000 millimeters which extends through more then half of the detector so that long proton and pion cluster like tracks will be found. The maximum start distance is set to be 250 millimeters which is about 5 modules within the same plane because the energy clusters from pions and protons should appear very near the vertex. The maximum gap distance is set to be 100 millimeters which is nearly 2 modules within the same plane because the energy clusters produced by the protons and pions should fall within this distance from one another.

Chapter 3

Neutron Identifying Algorithm

The study in this thesis requires the identification of neutrons as products from anti-neutrino interactions with hydrogen and carbon nuclei. This gives us another handle to show what improvements to the modeling better describe real reactions. The study can be done via scanning for neutrons by eye in the Arachne event display, which is thoroughly described in the first section of Chapter 2, but to get a large enough sample with high statistics this would require an enormous amount of time. The software algorithm I developed is a significant new reconstruction opportunity for this and future MINERvA analyses.

The process of writing the algorithm did begin with scanning. The MINERvA collaboration has a large sample of anti-neutrino Monte-Carlo (MC) simulated events. Since these events are expected to predict how particles behave in the MINERvA detector accurately they were used to help identify the differences between neutrons and other particles. To do this scanning 100 anti-neutrino events were selected at random from both the low and mid $q3$ regions. All of the information is available to the user for simulated events, including what particles were present, their kinetic energies, three-momentum, direction, re-interaction points, and the products of those re-interactions.

Using the information gathered from the events scanned several important facts about neutrons were confirmed and discovered. The neutrons from the anti-neutrino interaction do not always become visible in the detector by re-interacting and depositing energy. When neutrons do re-interact they typically do so away from the vertex and the muon track, as was expected due to their mean free path. During this process they typically deposit small amounts of energy in single scintillator strips ("hits") or in multiple nearby strips ("clusters"), which are visible in the Arachne event display. Typically they are few in number and also low energy, from elastic scatters with hydrogen and quasi-elastic proton knockout from carbon nuclei. Unfortunately these are aspects that can be true of other types of particles as well so there will be a background to these neutrons.

3.1 Algorithm Rules

The algorithm follows some very basic rules to identify possible neutron candidates by using information about clusters of energy (one or more "hits") deposited in the detector during each anti-neutrino interaction. Then the algorithm compares clusters that are near one another to in a way "combine" them to be counted as a single neutron candidate by creating something called a blob. A blob is simply combining multiple hits and clusters that are thought to be from one specific thing into one reconstructed object.

The rules of cluster selection and elimination begin with any cluster must have a kinetic energy equal to or above 1.5 MeV. There are many processes, noise or low energy photons, that leave very small amounts of energy in the detector, which I do not want to count as neutrons. While this does cause the efficiency of counting neutrons to decrease the loss is not significant enough to outweigh the gain in purity from the background elimination

The algorithm then eliminates any cluster that is within 10 planes vertically from the muon track. This distance was selected so the algorithm eliminated clusters deposited from photons, electrons, or positrons coming from the muon re-interactions that escaped the muon fuzz box originally used. Unfortunately there is no perfect distance that allows the algorithm to eliminate all products coming from the muon re-interactions which means these products typically leave small amounts of energy that make them appear very much like a neutron.

The clusters must then be more then 5 modules away horizontally and 10 planes vertically from the vertex. This helps eliminate clusters that come from energetic showers produced by high energy charged particles like protons and pions. However this does not eliminate all clusters coming from these types of particles so the algorithm also eliminates any cluster tagged to be on a track coming from the vertex or that is a part of a vertex blob. Both the muon fuzz box and the vertex blob tool are discussed in chapter 2.

Once the clusters are filtered and only the few that pass the first set of rules are left, the algorithm then loops through the list of clusters, starting with the farthest upstream and closest to the bottom of the detector while alternating between views, comparing the currently viewed one to all the other clusters already reviewed. This means the first cluster will always start one candidate and the second cluster in the list will be compared to only the first to either combine into a single candidate or begin a second. This then continues for the third cluster and so on.

Multiple clusters may be from a single particle, and my algorithm uses spatial proximity to group them into a single neutron candidate. Simply put the algorithm draws a small box around each cluster and combines any cluster or group of already combined clusters that have any fraction of itself within the box. The box is drawn to be 10 scintillator strips transverse (as reckoned in Arachne) and 3 modules longitudinally along the z axis of the detector. The algorithm then looks at clusters within different views but because of the difficult geometry of the detector it only considers the horizontal distance between clusters. If the most recent cluster being considered is within 3 modules, yet still in the different view, horizontally from a previously considered cluster or group of clusters they are combined into one candi-

date. Once the algorithm has looped through the entire list the final list of neutron candidates will then be available.

Data event neutron candidates have information saved about the deposited energy, starting and ending module, and transverse distance from the vertex. For simulated events all of the above is available along with all of the truth information, especially details about the simulated particle that caused the energy deposits.

In simulation each candidate counted by the algorithm is identified as being created by a specific particle. For each simulated event, every cluster is given something called a PDG code (after the Particle Data Group standard) that identifies what particle produced by GENIE (meaning it was a product that left the nucleus from the anti-neutrino interaction) caused that energy deposit. In the algorithm when multiple clusters are combined into one candidate the PDG code from the largest contributing cluster to the candidate kinetic energy is assigned to the candidate. While it does not happen often for a candidate to have more then one particle type contributing, it is possible. This information allows for a separation of the candidates into true signal, or GENIE neutrons, and background, which is everything else.

3.2 Tuning Parameters and Algorithm Studies

To test the parameters such as the lower energy cut to eliminate clusters created from background or the distances between clusters to form candidates the values were changed one at a time by increasing and decreasing them and then using the algorithm on MC events along with the truth information to get the best "efficiency" to find all the energy actually deposited by neutrons and to at least characterize or even reduce the backgrounds. The efficiency was then compared for each version of the algorithm to decide what is the most optimized version.

Neutron Energy	Low q3	Mid q3
Total (MeV)	236146	1629000
Algorithm Counted (MeV)	194422	1365630
Efficiency	82.33%	83.83%

Table 3.1: Algorithm's neutron energy counting efficiency

Before diving into the efficiency tables a few terms must be defined. First is a particle that is produced by GENIE: these particles are produced by the simulation and are said to have been the particle that left the nucleus whose constituents were involved in the anti-neutrino interaction. Geant4 is the particle physics standard simulation that steps particles through the detector to determine energy loss or interaction. It starts with the GENIE particles, though later re-interactions can produce additional particles, including neutrons. An example that helps illustrate this is as follows. GENIE produced a pi minus, meaning it left the hydrogen or carbon nucleus after the anti-neutrino interaction, which Geant4 interacted kicking out

a neutron which itself interacted or scattered off hydrogen, causing the recoil proton to make an energy deposit. This is a real neutron but it's not a GENIE neutron, so it doesn't directly tell us about the interaction model. The clusters deposited during these different interactions are assigned a PDG code for the parent particle, the GENIE particle leaving the nucleus, as well as one for the Geant4 particle. The PDG code from the GENIE particle is what is assigned to the candidate created by the neutron counting algorithm as this analysis is in search of describing the interaction model.

Table 3.1 gives the results for one way of illustrating the efficiency of the algorithm. This "energy counting efficiency" does not penalize the algorithm for not finding neutrons that weren't visible in the first place, because they left the detector without scattering. To create this table the energy of every cluster from each anti-neutrino event in the low q3 and mid q3 selections of the minerva5 playlist is summed if a cluster is tagged to be from a GENIE neutron. For this study the simulated 2p2h events are included and the RPA re-weight is being applied. Then the entire list of clusters is provided to the algorithm and the energy of every candidate tagged to be from a GENIE neutron is summed and compared to the previous value.

The algorithm's efficiency at counting neutron energy sits around 83 percent for both the low and mid q3. The majority of the loss of energy is likely due to the loss from the 1.5 MeV energy cluster cut put into the algorithm and from neutrons that interacted near the vertex. However elimination of the former cut decreases the purity of the sample quite significantly as background from photons and data overlay is picked up. It is important to not confuse the visible energy counting efficiency with the true energy of the neutrons. Studies discussing how much energy is visible compared to the true neutron energy and the overall neutron visibility can be found in later parts of this document.

Since the algorithm allows for the total number of candidates to be counted, as well as gives the starting and ending modules of each, the candidates can be separated by their position either behind or in front of the vertex. The motivation for this gives the user an idea of the range in angles of the neutrons produced. This can be done for both data and MC events and the true neutron angle is available for the later. Table 3.2 compares the total number of candidates per event, candidates in front of the vertex, and candidates behind the vertex between simulation and data for the QE and dip hadronic energy regions for the low q3. The mid q3 is not presented as the story is the same. The selection of where the candidate was in comparison to the vertex was done using the ending module. If the candidate module was equal to or higher then the vertex module it was in front of the vertex while anything else was considered behind.

It is clear from Table 3.2 there is a discrepancy between the data and the simulation. It appears while similar percentages of overall candidates are found to be moving forward and backward for the data and simulation, given either slice of momentum transfer, the total number of candidates per event is very different. The simulation has a significantly higher number of neutron candidates per event then the data does which suggests the simulation is not assessing the cross section for neutrons, and potentially other particles, accurately. This agrees with many more

QE Region

Simulation	Percent Forward	Percent Backward	Data	Percent Forward	Percent Backward
Algorithm/ Uncertainty	67.31 % 0.25 %	32.69 % 0.25%	Algorithm/ Uncertainty	66.50 % 0.78 %	33.50 % 0.78%
Candidates Per Event	0.30 ± .006	0.15 ± .004	Candidates Per Event	0.29 ± .020	0.15 ±.015

Dip Region

Simulation	Percent Forward	Percent Backward	Data	Percent Forward	Percent Backward
Algorithm/ Uncertainty	64.93 % 0.74%	35.07 % 0.74%	Algorithm/ Uncertainty	61.32 % 1.70 %	38.68 % 1.70%
Candidates Per Event	0.74±.012	0.45±.009	Candidates Per Event	0.57±.028	0.37±.022

Table 3.2: This table shows the percentage and associated statistical error of candidates moving forward or backward for simulation and data in the QE an dip hadronic energy regions. The total number of candidates per event moving forward and backward along with the associated statistical error for each of these regions is shown as well.

of the studies presented in further portions of this document.

Neutrons frequently do not interact before they leave the detector. The simulation gives us an estimate of how visible they are using truth information along with the algorithm. Selecting events that have one neutron (not necessarily a visible neutron) and no other particle type, separating the events by how energetic the neutron is, and then counting neutrons with the algorithm provides how often a neutron interacts as a function of their energy. It is ideal to remember given a neutron was present and visible the algorithm is capable of counting approximately 83 percent of that energy. This also shows how often the algorithm may count more then one candidate for a single neutron if it scatters more than once or is energetic enough to produce clusters distant enough from one another.

Figure 3.1 shows the information described above. From the first plot it is apparent neutrons with kinetic energy of 10 MeV or less will almost never interact in the detector and become visible. The second plot can be interpreted to mean neutrons above 10 MeV and below 50 MeV will be visible approximately 30 percent of the time while sometimes, but not often, re-interacting more than once causing more then one candidate to be counted. The third plot follows the same trend as the second with the visibility of the neutrons increasing to nearly 45 percent with the energy increasing to be between 50MeV and 100 MeV and the re-interactions increasing as well. The story remains the same for the final two plots with even higher ranges of neutron kinetic energy. The number of neutrons causing two candidates to be counted is something that can be changed by altering the distance parameters

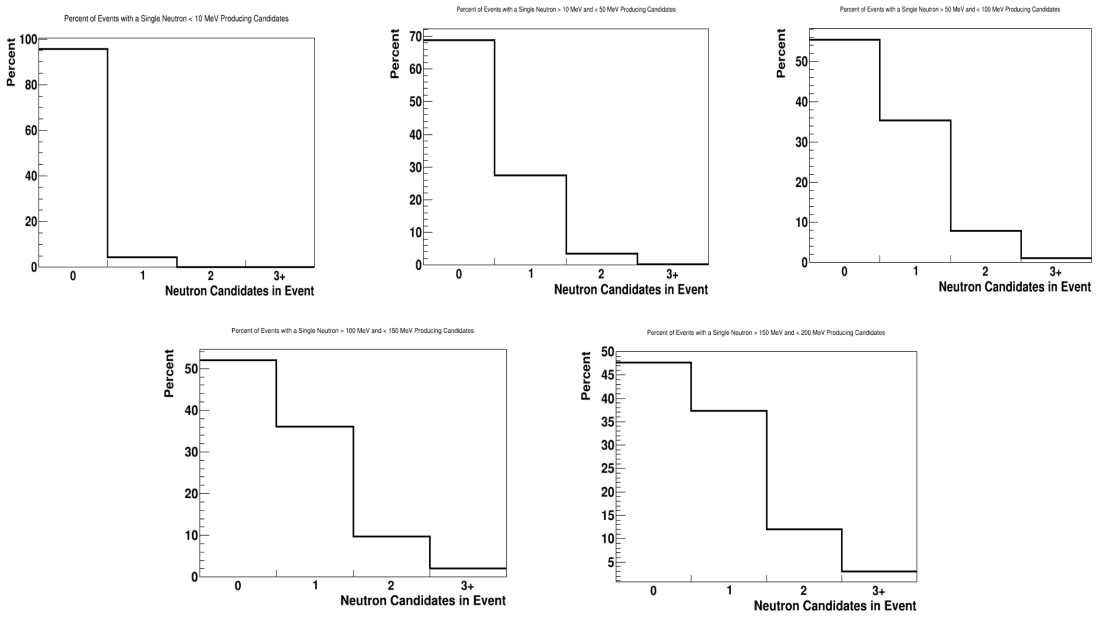

Figure 3.1: These plots show the total number of candidates counted per event with only a single neutron with: less then 10 MeV (top left), 10 -50 MeV (top middle), 50-100MeV (top right), 100-150MeV (bottom left), 150-200MeV (bottom right). Both low and mid q3 were considered when making this plot. The lowest energy neutrons are nearly invisible as we see them less then 10 percent of the time

between clusters to be considered one single candidate. However increasing this parameter would cause a loss in accurately predicting number of neutrons per event when there is more then one present.

3.3 Backgrounds

The optimized version of the algorithm does not provide a completely pure sample of candidates either. The number of candidates contributed by each particle type can be accessed using the GENIE PDG codes, because we want to know what GENIE particle caused this not what Geant4 particle did, when looking at MC events. Table 3.3 shows exactly what the signal and background is for the entire sample of events in both the low and mid q3 regions. The low q3 candidates are primarily neutrons with the second largest contributor being π^-. The mid q3 candidates follow the same trend but pi minuses now become a larger fraction of the total and neutral pions also become more prevalent. Neutral pions are a possible reducible background given they typically emit a photon when they undergo secondary interactions which could possibly be an easy π^0 indicator. However in the next few paragraphs it will become clear that π^-, neutral pions, and protons are all irreducible backgrounds for the algorithm.

Knowing the types of background the algorithm counts prompts several other studies that include determining whether these backgrounds are irreducible or not. The first study includes scanning events by eye with a single GENIE pi minus or

Particle Type		Counted Candidates	Percent of Total
Neutrons		22908	76.60%
Protons		83	00.28%
Pi Zero		806	02.70%
Pi Minus		3228	10.79%
Pi Plus		5	00.02%
Photons		198	00.66%
Data Overlay		857	02.87%
Electron/ Positron		0	0.00%
Muon		1817	06.08%
Other		3	00.01%
Total		29905	100.0%

Particle Type		Counted Candidates	Percent of Total
Neutrons		81034	59.86%
Protons		2073	01.54%
Pi Zero		14298	10.56%
Pi Minus		30197	22.31%
Pi Plus		232	00.17%
Photons		664	00.49%
Data Overlay		2670	01.97%
Electron/ Positron		0	00.00%
Muon		4119	03.04%
Other		75	00.06%
Total		135362	100.0%

Table 3.3: Candidates counted by algorithm separated by particle type using truth information. The low q3 is on the left and the mid q3 is on the right

a single GENIE proton to see how the clusters coming from these particles are not being eliminated and are counted as a candidate. Using truth information such as the PDG codes, module, and view along with Arachne shows that the clusters being tagged are actually particles that are products of secondary interactions.

Other	11/20
Pion induced neutron	6/20
Pion	3/20

Other	8/20
Pion induced neutron	9/20
Pion	3/20

Table 3.4: The table shows what product particle produces a cluster from a GENIE pi minus undergoing a secondary interaction. The low q3 is on the left and the mid q3 is on the right

Tables 3.4 and 3.5 are for events with either a single proton of a specific energy range or a pi minus. For the pi minus case the table shows the counted candidates are typically neutrons or other low energy particles such as photons that come from the pi minus undergoing secondary interactions. Unfortunately low energy photons produce energy clusters similarly to neutrons and while we do want to count neutrons we do not want the products from other particles interacting. These are irreducible backgrounds. The proton case is very similar in that almost all of the clusters counted by the algorithm are neutrons that came from secondary proton interactions. This case was split up by the energy of the proton to determine if protons behaved in different ways as a function of their energy. While there is a small dependence between proton energy and the overall momentum transfer there

Energy	< 20 MeV	20 MeV < E
Proton induced neutron	18/20	18/20
Protons	2/20	2/20

Energy	< 20 MeV	20 MeV < E
Proton induced neutron	19/20	17/20
Protons	1/20	3/20

Table 3.5: The table shows what Geant4 particle produces a cluster from a GENIE proton undergoing a secondary interaction. The low q3 is on the left and the mid q3 is on the right

is nothing useful to use to eliminate extra background.

Knowing that so many candidates are being contributed by pi minuses also prompts a study to determine the efficiency of the algorithm when counting pi minus's. As previously discussed, the truth information provides the pdg codes for each cluster. Table 3.6 shows the energy summed for every cluster in each event tagged to be from a GENIE pi minus compared to the energy summed of all the candidates tagged to be from a GENIE pi minus created by the algorithm.

Pion Energy	Low q3	Mid q3
Total (MeV)	58781	1091840
Algorithm Counted (MeV)	32565	588079
Efficiency	55.40%	53.86%

Table 3.6: The table shows the algorithm's efficiency at counting pi minuses by comparing total pi minus visible energy to the algorithm counted pi minus energy

Table 3.6 helps show how much energy is both eliminated by the algorithm and also contributes to the neutron candidate energy provided by pi minuses. The result is the algorithm can eliminate approximately half of the pi minus energy available in low q3 and mid q3 region events. The half that is eliminated can be contributed to the algorithm's ability to detect unwanted clusters caused by particles that interact near the vertex such as pi minuses. Unfortunately the other half of the energy that is counted is due to the secondary interactions noted in the previously mentioned study. Even though it is irreducible, it is useful to know approximately how much of the total energy counted by the algorithm is due to this particular background.

Unlike protons and pions which create an irreducible signal that is actually neutrons from Geant4, neutral pions produce photons and never neutrons. This more sophisticated study was completed to determine whether or not the π^0 background found in the mid q3 is irreducible. To accomplish this approximately 100 MC events in the mid q3 with at least one π^0 were scanned by eye. The events in consideration also had to have at least one neutron candidate counted by the algorithm that was considered to be from a π^0. In these events the π^0 candidates were carefully examined to determine how the interaction appeared in the MINERvA detector. The mid q3 was chosen due to its large pi zero background content. Neutral pions do appear

in the low q3 but are much less of a problem and are expected to be of lower energy, and as us will see, causing them to appear more often in ways indistinguishable from neutrons.

Fortunately neutral pions interact in only a few ways. The π^0 decays 99% of the time to two photons, sometimes doing so before leaving the nucleus where it was produced. What these photons do help determine whether or not a π^0 was present as well as what type of tool can be used to help eliminate them from the background. Typically the energy trail left behind from the π^0 and one of the photons lies on a straight line coming from the vertex of the reaction. Due to their mean free path it is common for the photons to travel a distance before interacting and depositing energy within the detector. From the truth lines available in Arachne is it possible to see that these clusters of energy do come from the photons produced in the π^0 decay.

The π^0 can also produce photons that leave energy showers that are large and have the characteristic of reaching their max energy deposit in the middle of the shower before slowly decreasing and eventually dying out. The second photon in either of these reactions usually is of much lower energy and can either be invisible by not re-interacting or produce low energy clusters that appear very similar to low energy neutrons. Examples of these interaction types can be seen in Figure 3.2 and 3.3.

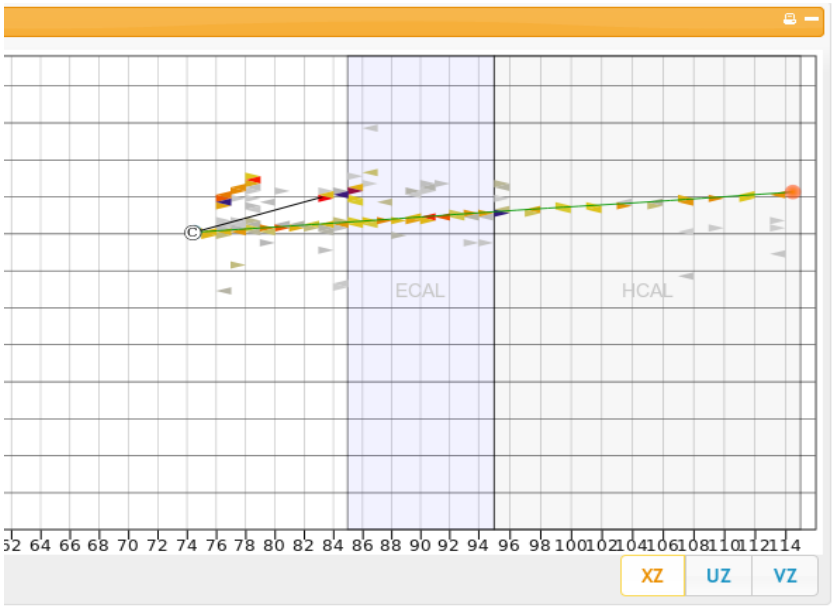

Figure 3.2: Event from MC run-50201, subrun-99, and gate-942. The event includes a neutral pion producing two photons who in turn interact in the detector. Both producing energy showers and one whose maximum deposit appears in the center of the shower. Darker colored areas correspond to higher energy deposits

These higher energy photons can still produce energy clusters in similar ways as neutrons do such as a single cluster of energy from 10 MeV to 30 MeV. This would put events with a π^0 that only produces photons that interact in a similar way to a neutron, meaning their interactions are distant from the vertex by their

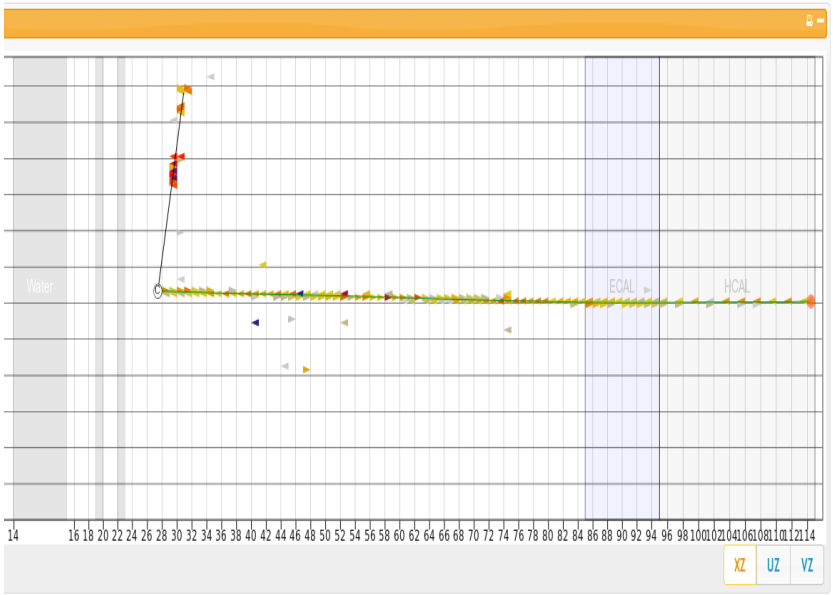

Figure 3.3: Event from MC run-50201, subrun-91, and gate-643. The event includes a neutral pion producing two photons. One is invisible to the detector and the other produces several clusters as it interacts, producing electrons, following along a straight path.

radiation length and produce single clusters, in the irreducible background category. The other types of events can possibly be eliminated using several tools. Existing already are a π^0 tool, which searches for an electromagnetic shower with the energy distribution peaking in the middle, and a tool that searches for energy deposits that lie within the same few degrees but at different radii's coming from the vertex, sort of creating a cone extending from the vertex. 100 MC events were scanned from the mid q3 containing neutron candidates produced by neutral pions. These candidates were then separated into three categories depending on whether it could be eliminated via the π^0 electromagnetic shower tool, the cone tool, or if it was too neutron-like to be eliminated.

Pi Zeros

Tool	No Fix (Neutron Like)	Eliminated by Cone Tool	Eliminated by Shower / Energy Distribution Tool
Percentage of Candidates	40.95%	39.05%	20.00%

Table 3.7: Results of scanning events in the mid q3 that contain a neutral pion producing a neutron candidate counted by the algorithm. Candidates are placed into categories corresponding to tools that could be used to eliminate them. Almost half can not be eliminated by the tools.

The results from the scanning are shown in table 3.7 and it is apparent that a large fraction of the π^0 contribution to the neutron candidate population is irre-

ducible. Many candidates that could be eliminated via the energy distribution tool could also be eliminated via the cone tool and were placed in the later so that only one category for each was chosen.

While knowing whether or not the neutral pions can be eliminated from background is important it is also important to understand how the neutron candidates or the signal will be affected. Two different studies were done to try to understand this without going through the lengthy job of implementing either of these tools and then looking at the new signal and background. The first was similar to the π^0 scanning just discussed. Instead of scanning for events with a π^0 producing a neutron candidate, events from the mid q3 with a single neutron that has 50 MeV of kinetic energy and produced a candidate were selected. The 50 MeV of kinetic energy was chosen to increase the chances of having a visible neutron in the event and to increase the probability of it depositing more energy then of a lower energy neutron. The events were then separated into the same three categories as before to determine how often these candidates would also be eliminated by the tools if they were used.

Neutrons > 50 MeV

Tool	No Fix (Neutron Like)	Eliminated by Cone Tool	Eliminated by Shower / Energy Distribution Tool
Percentage of Candidates	57.00%	23.00%	20.00%

Table 3.8: Results of scanning events from the mid q3 sample with a single neutron with kinetic energy greater then 50 MeV that was counted as a neutron candidate. Candidates are placed into categories corresponding to tools that would eliminate them. More then half can't be eliminated by a tool.

The results shown in table 3.8 demonstrate that higher energy neutrons can produce energy distributions in a similar way to neutral pions. From the truth information provided by GENIE and Geant4 it is possible to grasp at a few reasons why this is. Using the particle parentage history available for simulated events it is possible to see where each parent particle resulting directly from the anti-neutrino interaction goes as well as what it produces and where during the re-interactions along this path. These trees indicate that the higher energy neutrons are typically producing photons directly in their interactions with nuclei, or producing neutral pions which decay to photons as already described. This unfortunately means the higher energy neutrons look almost exactly like neutral pions to the MINERvA detector.

The final questions then are how many of the neutron candidates coming from neutrons of any energy would be eliminated using these tools and is it worth losing this signal if a large fraction of the background is also eliminated? The second π^0 study completed gives an approximation to the population of the signal candidates and background candidates eliminated using one of the π^0 tools mentioned. A similarity between the candidates in the events that could be eliminated by these

tools are that they are of higher energy which directly corresponds to the number of modules the particle distributes energy in next to one another. The algorithm returns the number of modules spanned by the candidate. This allows for candidates that have an energy deposit spanning any given number of modules to be eliminated. Plotting the number of candidates from both background and signal versus their module span leads to whether or not there is a module span number that eliminates significant background but leaves the signal untouched.

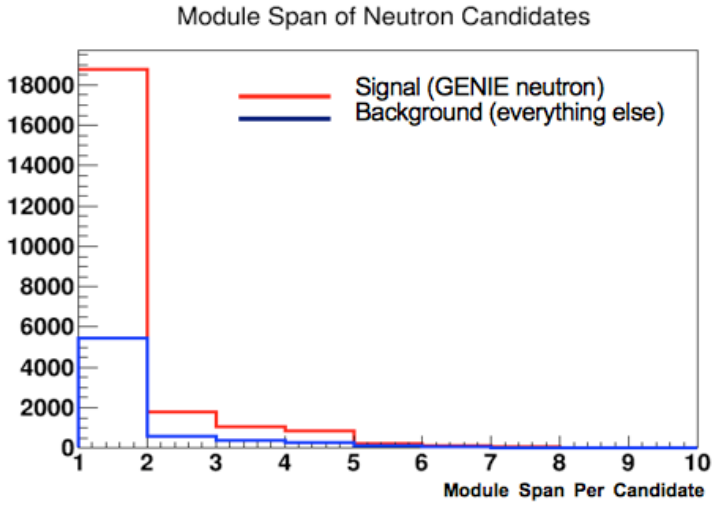

Figure 3.4: The plot represents the module span of both background and signal candidates found in the entire low q3 sample. It is clear there is not a number of modules a candidate can span that can be used as a cut in the algorithm to eliminate background and ignore the signal.

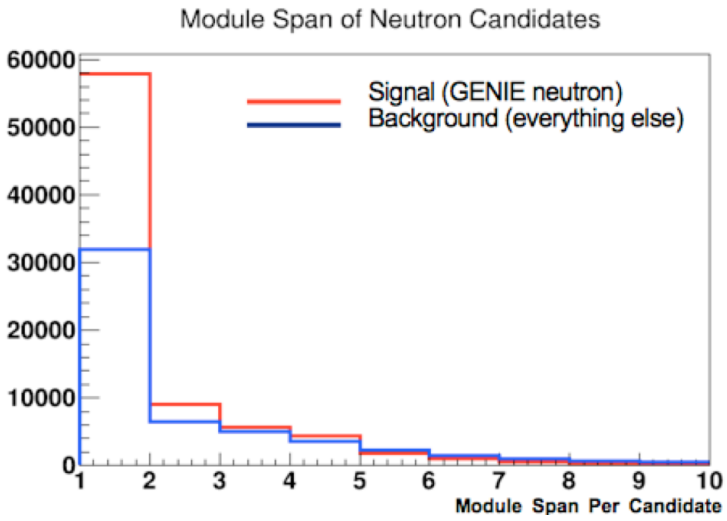

Figure 3.5: The plot show the module span of both background and signal candidates found in the mid q3 sample. The neutrons and backgrounds at higher momentum transfer are also necessarily higher energy on average, and their products produce larger and longer energy deposits. A selection to keep events shorter than 5 (or maybe 3) modules could increase purity in the mid q3 sample, but it wouldn't be a huge win.

Low q3

Particle Type		Counted Candidates	Percent of Total
Neutrons		22475	76.97%
Protons		82	00.28%
Pi Zero		679	02.33%
Pi Minus		3115	10.67%
Pi Plus		5	00.02%
Photons		195	00.67%
Data Overlay		833	02.85%
Electron/ Positron		0	0.00%
Muon		1809	06.20%
Other		3	00.01%
Total		29196	100.0%

Mid q3

Particle Type		Counted Candidates	Percent of Total
Neutrons		76968	61.32%
Protons		1992	01.59%
Pi Zero		10847	08.64%
Pi Minus		26877	21.41%
Pi Plus		200	00.16%
Photons		1992	01.59%
Data Overlay		2482	01.98%
Electron/ Positron		0	00.00%
Muon		4090	03.26%
Other		63	00.05%
Total		125511	100.0%

Table 3.9: The table shown here represents the background and signal after utilizing a cut in the neutron counting algorithm that ignores candidates who span over 5 or more modules. The low q3 (left) and the mid q3 (right) are shown. The background percentage does not change enough to make the cut useful.

The plots shown in figure 3.4 and 3.5 help demonstrate that a specific module span for neither the low q3 nor mid q3 can be selected to help maximize the signal

and eliminate the background. This correlates to neither of the tools mentioned being useful enough to warrant the time needed to implement them into the current algorithm. However for completeness the algorithm was adjusted to eliminate candidates that spanned over 5 or more modules and the total background signal and remaining can be seen in Table 3.9.

Comparing table 3.9 with the previous background counts in table 3.3 helps complete this story. Both the low and mid q3 lose background and signal due to this new cut being implemented. For the low q3 the total background is reduced by 5.26 percent, the total signal is reduced by 3.03 percent, and the new fractions of background and signal candidates compared to the total number of candidates are nearly the same as they were before. In the mid q3 the total background is reduced by 16.66 percent, the total signal is reduced by 9.9 percent, and the new fractions of background and signal candidates follow the same trend as the low q3. This means while the tools could potentially eliminate background it will also eliminate enough signal so that the overall candidate sample looks the same but is now smaller.

3.4 Neutron Visible and True Energy

There is no correlation between the amount of energy a neutron has and the amount of energy a neutron deposits in the MINERvA detector. Figure 3.6 contains the plots of the visible energy from the neutron candidate created in an event that had only a single neutron and no other hadrons versus the truth energy of the neutron in that event. The plots show that no matter what the true energy of the neutron in the event was, the candidate had 10 MeV of energy or less most of the time. This is easily seen if looking at the bins horizontal to the arrows for each plot. It is clear the first two bins, corresponding to 10 MeV or less, in this direction contain most of the candidates.

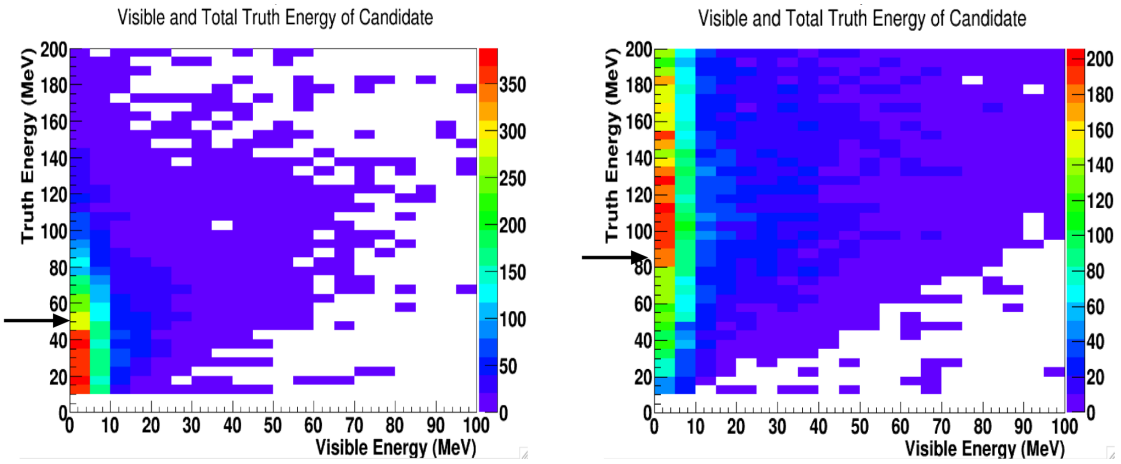

Figure 3.6: Visible energy of a neutron candidate plotted against the truth energy of the neutron in an event for low q3 (left) and mid q3 (right). Events were selected to have a single neutron and no other hadrons so neutron candidate had to be from the single neutron.

3.5 Algorithm Summary

As a quick overview the main features of the candidates selected by the algorithm are listed here. The candidates are created from clusters not in the filament vertex blob, they must be at least 100 millimeters away. These clusters are combined spatially into candidates, with an energy, position information. From the neutral pion studies, if it spans 5 or more modules, the candidate was eliminated.

Given the examples presented through out this chapter it is obvious the algorithm performs fairly well. The overall efficiency of the algorithm's ability to count visible neutrons is around 83 percent while the purity of the sample of candidates obtained by it are between 60 and 75 percent. The studies discussed here show the algorithm has been tested for possible improvements in a variety of ways and is in its most optimized version. An example of the neutron multiplicity plots created using the algorithm that are used throughout the rest of this paper is shown in Figure 3.7.

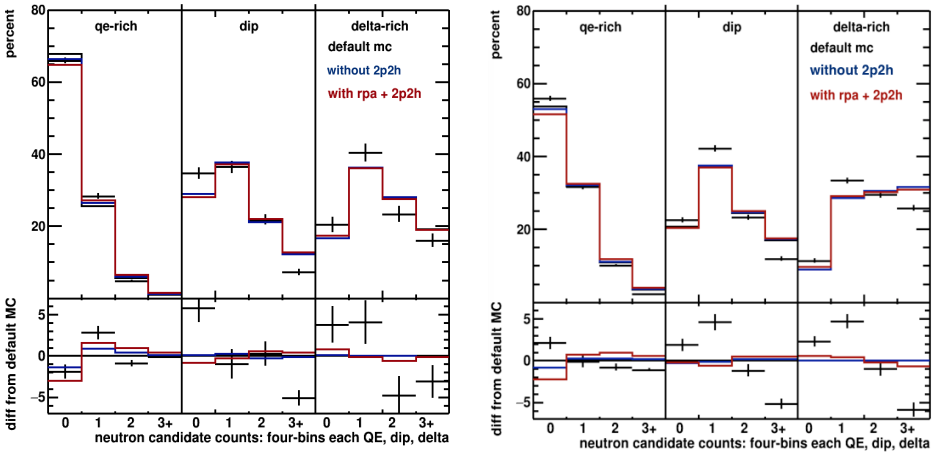

Figure 3.7: Neutron counting plots (middle) include data, the default simulation, default plus RPA, and RPA plus 2p2h. Plots for the low q3 (left) and mid q3 (right).

The neutron counting plots can be difficult to understand so an explanation is deserved. The top portion of the neutron counting plots show the percentage of events with 0,1,2, or 3 plus neutron candidates using the data and simulations described for each of the regions of available energy determined by the QE, delta, and 2p2h interactions. Remember the QE hadronic energy region lies between 0.0 and 0.07 (0.09) MeV , the dip region then between 0.07 (0.09)MeV and 0.14 (0.21) MeV, and the delta region from 0.14 (0.21) MeV and up for the low q3 (mid q3). The bottom portion of the plot shows the difference between the data and each simulation type compared to the default simulation. This portion allows for a quick determination of what matches the data the most.

The event selection used for the studies discussed in this section and the rest of this paper was made following the prescription described next. It is an inclusive sample from the minerva5 anti-neutrino playlist where there is 1.02177e20 data POT and 9.25444e20 MC POT, the reconstructed anti-neutrino energy must be between 2 and 6 GeV. This includes MINOS match and the apothem fiducial cut with the

muon angle being less the 20 degrees with respect to the direction of the beam. MC gets the standard cv weights. The RPA and 2p2h here are not of the newest forms. The available energy is only tracker and ECAL, passive material corrected only. The muon fuzz, all non-muon energy in a cylinder around muon, is subtracted. The available energy is scaled by 1.0 as a replacement for Anne's poly-line/spline correction.

Chapter 4

Neutron Position Distributions

Chapter 4 will develop observed neutron-candidate position distributions: downstream to upstream (z-distribution) and transverse. This is complementary to Chapter 3 which introduced the observed neutron-candidate relative multiplicity distributions with an initial comparison of data and simulation. Then Chapter 5 will combine the two and systematically compare different model elements, RPA and 2p2h, and systematic uncertainties in the MC to the data with both the z-distributions and neutron multiplicity distributions side by side.

4.1 True Angle Emission for Interaction Processes

The angle and position of neutron candidates were studied to understand the differences among the QE, 2p2h, and Delta interaction models, as well as the signal and background. To study what affects the spatial distribution of neutrons, we used a subset of the MC sample that had only a single neutron and no other hadrons. Most of these events fall in the QE hadronic energy range because that process is the most likely to produce a single neutron product from the interaction. This selection means if a candidate was counted it most likely came from the neutron in the event and thus it was visible. Using conservation laws it is expected for the neutron in a pure quasi-elastic interaction to exit the nucleus near a 90 degree angle from the outgoing muon. At these beam energies and momentum transfers, this is almost always a little less than 90 degrees from the Z axis. However, for the MINERvA detector, exiting the nucleus at a 90 degree angle typically allows for the least amount of traveling distance before leaving the detector. Since neutrons have a long interaction path length this causes for events with visible neutrons to not peak at a 90 degree neutron angle. This is demonstrated in Figure 4.1. The peak is shifted slightly lower to be just below 80 degrees for both the low and mid q3 plots. It is also very noticeable that very few candidates are initially moving behind the vertex, having angles greater than 90 degrees.

The approximate peak of the neutron angles for the delta and 2p2h interactions should not be at 90 degrees due to conservation of momentum, because the hadronic momentum is shared with at least one other particle. These peaks were approximated by making the same plots shown in Figure 4.1 but with a selection

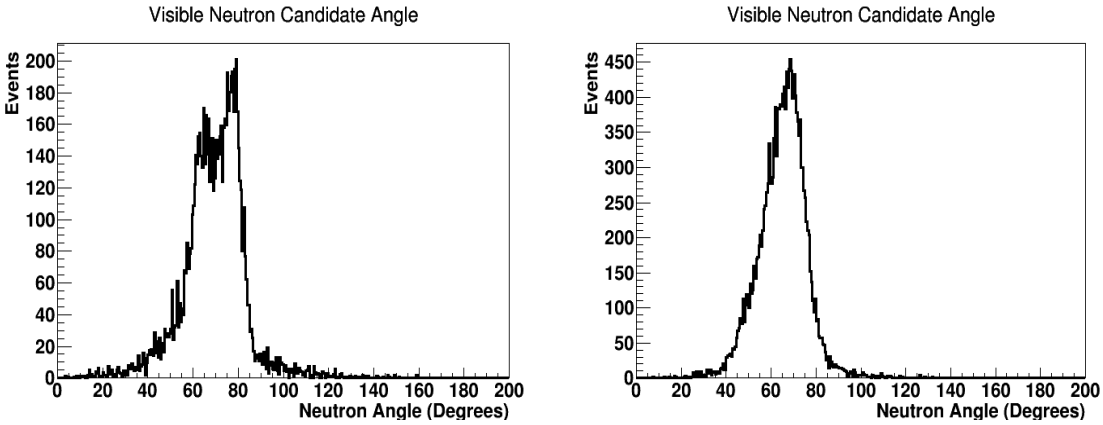

Figure 4.1: Truth angle of neutron in simulated events from the low q3 (left) and mid q3 (right) selected to have a single neutron and no other hadrons. Peaks are below 90 degrees

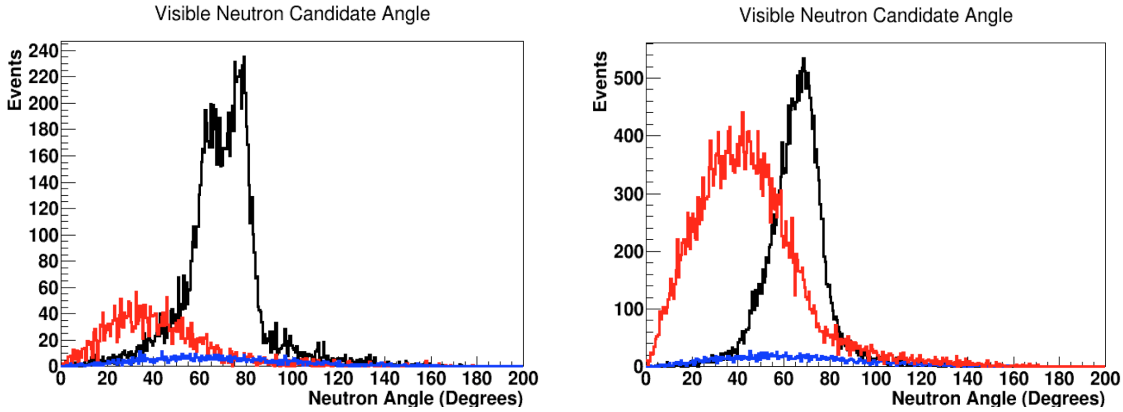

Figure 4.2: Truth angle of neutron in simulated events from the low q3 (left) and mid q3 (right) selected to have a single neutron and anything else. Plotted by interaction types where QE (black), delta (red), and 2p2h (blue). 2p2h and Delta have more forward moving neutrons.

of events with a single neutron plus anything else which are shown in Figure 4.2. The candidates here could potentially be from the protons and pions visible in these interactions but, given the pion caused neutron candidate background is not large and the proton background is very small, most should be from the neutron. The peaks from the interactions are less pronounced and indicate the candidates could be found anywhere forward of the interaction vertex but it is still true that few move behind the vertex initially.

4.2 True Emission Angle and Kinetic Energy Relationship

We expect a correlation between the neutron energy and its angle, from conservation of energy and momentum. The least energetic neutrons, most likely from quasi-elastic events, are the most transverse. As a sanity check the link between the kinetic energy of the neutron and its angle with respect to the interaction vertex was made from the MC. Events with a single neutron were selected and the angle of the neutron was plotted against its kinetic energy. In Figure 4.3 there is a noticeable downward slope of 15 degrees in neutron angle per 200 MeV neutron kinetic energy range. This makes sense when considering conservation of momentum during the collision of the anti-neutrino with a nucleon along with the products afterwards. The more energetic neutrons must carry more of the original neutrino momentum, and so need to move in a more forward motion to conserve momentum in both the vertical and horizontal directions, assuming the anti muon from the interaction is also moving nearly forward.

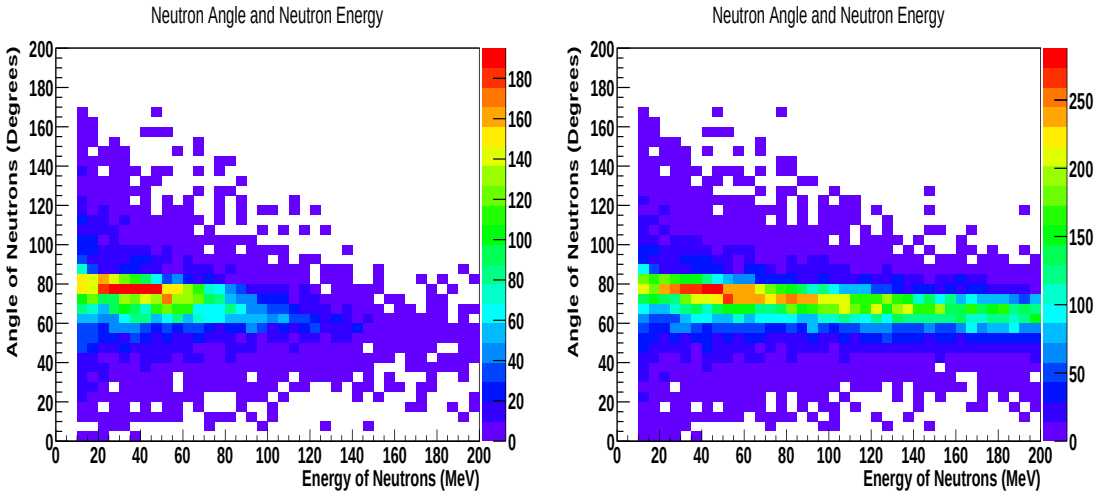

Figure 4.3: Plots of the neutron angle versus the kinetic energy of the neutron for the low q3 (left) and the mid q3 (right). Plot was made selecting MC events with only a single neutron and nothing else. The downward slope visible across the two plots is the correlation between energy and angle.

4.3 Neutron Visibility Dependence Upon Neutron Angle

The visibility of neutrons does not depend strongly on its emission angle, despite the geometry of the detector. This is in contrast to a strong dependence on energy. Again, the single neutron along with nothing else event sample was used to study this. The events were also selected to have a neutron with an angle within a specific range, where the angle is given from truth information, and then the number of

candidates in each event were plotted. Given there is only one neutron in each event there should typically be either zero or one candidate per event with two occurring rarely when the neutron was energetic enough to cause multiple re-interactions.

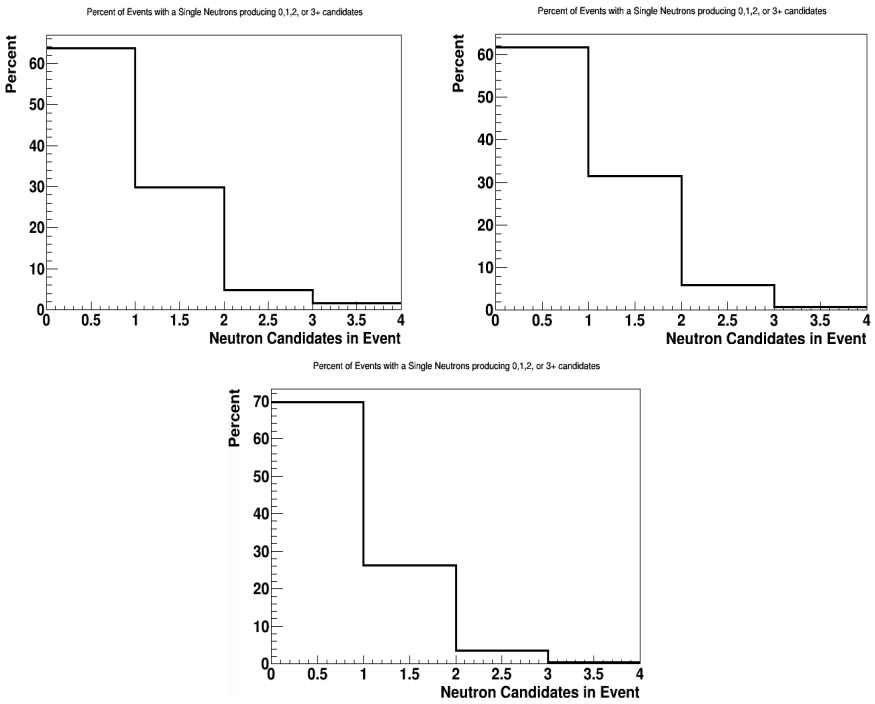

Figure 4.4: Percent of events that have 0, 1, 2, or 3 plus candidates per event. Events were selected to be from both low q3 and mid q3, have a single neutron plus no other hadrons, and have a neutron angle relative to the direction of the exiting muon of 0 to 40 degrees (left), 40-80 degrees (middle), or 80-120 degrees (right). Forward moving neutrons are more visible.

Neutrons with angles between 0-40 degrees and 40-80 degrees have approximately the same visibility as one another with about 38 percent of the events in the low q3 and mid q3 having one or two neutron candidates found. However neutrons with angles between 80-120 degrees are visible only 30 percent of the time. When the neutron is not extremely transverse to the beam, neutron visibility is not strongly dependent on angle. This decrease is due to the shorter amount of detector available for the neutron to interact in when the neutron moves directly above or below the vertex. It would be likely if this table was broken down into smaller sections the least visible neutrons would have an angle around 90 degrees. However there is not sufficient MC events to create reliable plots at these angles.

4.4 Observed Candidate Z-Position from Neutrons with Peak Angles

The neutron candidates caused by neutrons with an angle inside the peaks found in Figure 4.1 could potentially scatter once without leaving a candidate, and then

move in any direction. They could also be made up of something other then just quasi-elastic events if the selection is changed to be events with a single neutron plus anything else as seen in 4.2. Neutron candidates are primarily observed in the slightly forward direction, but there are significant tails far forward and backward, when the candidate's z-distribution from the interaction vertex is plotted. This z-distribution is shown in Figure 4.5 for neutrons in two different ranges of angle, those ranges being within the peak for either the low q3 or mid q3 shown in plots above. The z-distribution is shown in number of modules in the detector where the closest module the candidate interacts in to the vertex is used. This would still be dominated by QE like events because these would only have a single neutron in the event. However there are 2p2h and delta interactions that also have a single neutron in them and so these also show up.

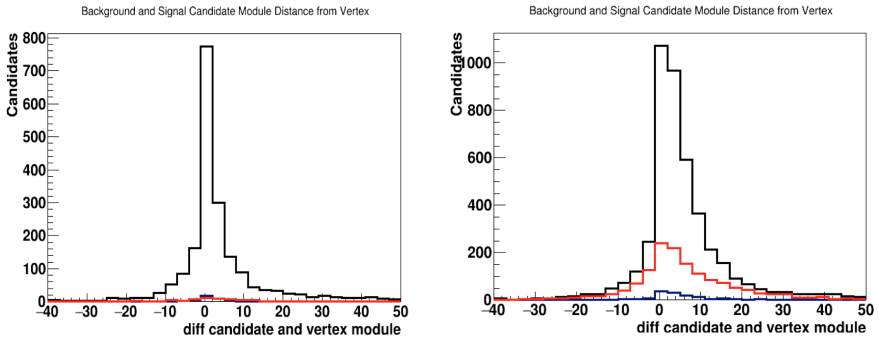

Figure 4.5: Plot of the difference between the candidate module and the interaction vertex module for low q3 (left) and mid q3 (right). MC events are separated by interaction type. Events were chosen to have a single neutron with a neutron angle between 75-80 degrees (left) and 65-70 degrees (right). Black is QE, red is Delta, and blue is 2p2h.

These distributions have a pronounced peak in the bin within five modules of the interaction vertex. This is indicative of the neutron candidates moving somewhat along the path of its angle from the interaction vertex but with some deviation. When comparing candidates based on what interaction type they were involved in, it seems as though events with a single neutron have a more pronounced peak if they came from a QE event. The peaks for the 2p2h and delta events are more broad and show the neutrons tend to travel more within the detector. This could be because neutrons from the 2p2h and delta interactions are expected to be more energetic. The more energetic the neutron is the more energetic the products of its own interactions could be. It is not expected for the neutron to leave with an angle near 90 degrees for these interactions as it is for the quasi-elastic.

4.5 Full Z-Distribution of Simulation in Comparison to Data by Process and by Signal and Background.

The distributions of the location of neutron candidates in the MC can be compared to data, and we begin to consider what aspects of the model might need the most adjustment. The simulation considered here always has RPA and 2p2h turned on with an additional systematic as well when noted. The addition of these extra systematics can help identify the mentioned aspects of the model that need adjustments.

Because the neutrino momentum is forward, and so is the momentum transfer to hadrons, the simulation expects the neutrons to exit the nucleus moving forward and very few events with a neutron moving backward (angles greater than 90 degrees). However, the Figures 4.6 and 4.7 show many more neutron candidates in the backward direction. These histograms were made by plotting the number of modules in between the first module upstream (nearest where the beam enters the detector) of each candidate and the vertex module, or first module downstream if in front of the vertex, for data and simulation with the simulation broken up by interaction type. The bottom of these plots show the ratio of the data and the simulation with RPA and 2p2h to the default simulation which has no additional weights incorporated. The module distance plots show the peak occurring in the middle bin. This would be indicative of the candidates moving just above or below the vertex. However, this bin spans several module differences meaning the candidate could technically have been at nearly any angle from the vertex depending on how far it is vertically from vertex itself and still within this bin.

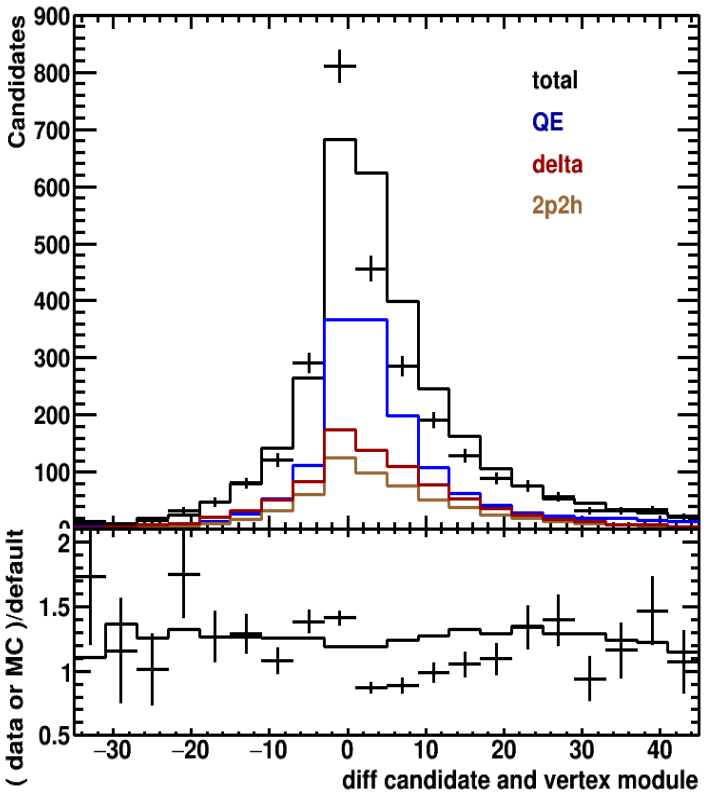

Figure 4.6: Plot of the difference between the candidate module and the interaction vertex module for low q3 MC events separated by truth event type. Data points with corresponding statistical error bars are plotted over top. Blue is QE, red is Delta, brown is 2p2h, and black is total. There is an over simulation of forward candidates.

The module distance plots show nearly one-third of the candidates in both the low q3 and mid q3 are found behind the interaction vertex, though the rate of those is well simulated. In the forward direction, the simulation over predicts candidates by 20-30 percent in each bin when 2p2h and RPA are included in the model. From scanning of MC events in Arachne and using the truth information on where the neutron moved throughout the detector it is apparent many neutrons interact after exiting the nucleus, often without depositing energy in the detector but changing direction and/or creating a new particle, which sometimes move backwards. These new particles then move far enough to be behind the vertex before interacting themselves and depositing visible energy. While these new particles may not be neutrons they hold the energy of the GENIE neutron so it's good that we count them. This encompasses all of the backward moving candidates counted.

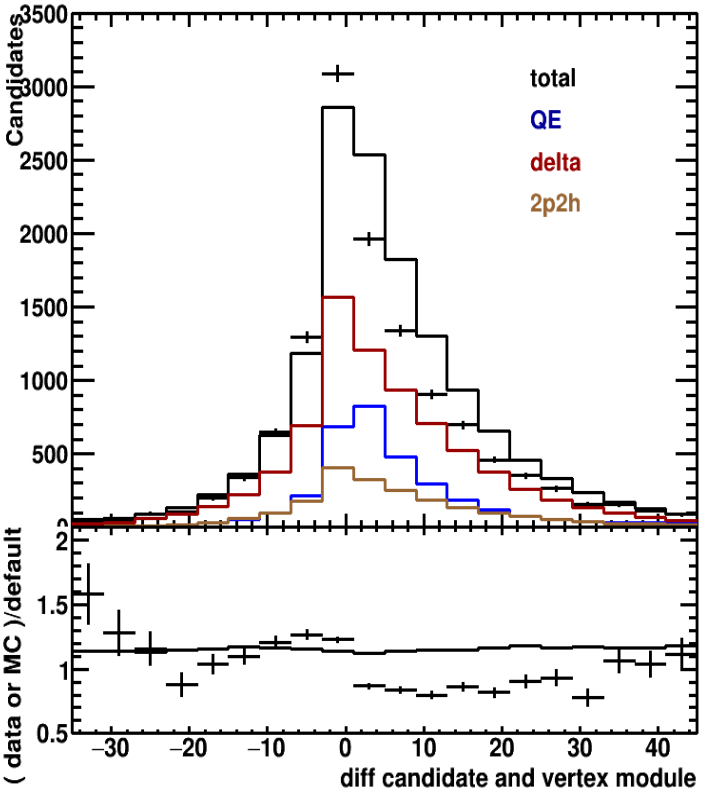

Figure 4.7: Plot of the difference between the candidate module and the interaction vertex module for mid q3 MC events separated by truth event type. Data points with corresponding statistical error bars are plotted over top. Blue is QE, red is Delta, and brown is 2p2h, black is total. There is an over simulation of forward candidates.

The total overall simulated candidates moving backward tends to agree with data for both the low q3 and mid q3 plots but this is not true for the forward moving candidates. Here we are greatly over simulating the forward moving candidates and thus also the total number of candidates as well.

Unfortunately there is little difference between a candidate that came from one interaction type or another. All three of the interaction types being modeled peak in the first bin for these histograms whether you are looking at the low or the mid q3. The quasi-elastic interactions are only slightly less forward than the others. In Figure 4.6 one can notice that no interaction type really causes a noticeable discrepancy between simulation and data. However, when looking at Figure 4.7 for the mid q3 it appears as though the delta interactions almost match data without the quasi-elastic or 2p2h interactions being added. This does not mean there is not a problem with the simulation of the neutrons causing neutron candidates in the quasi-elastic or 2p2h interactions but it does indicate there could be a problem in the delta type interactions.

It is possible neutron caused candidates appear in different areas of the detector compared to neutron candidates caused by other particles leaving the nucleus in

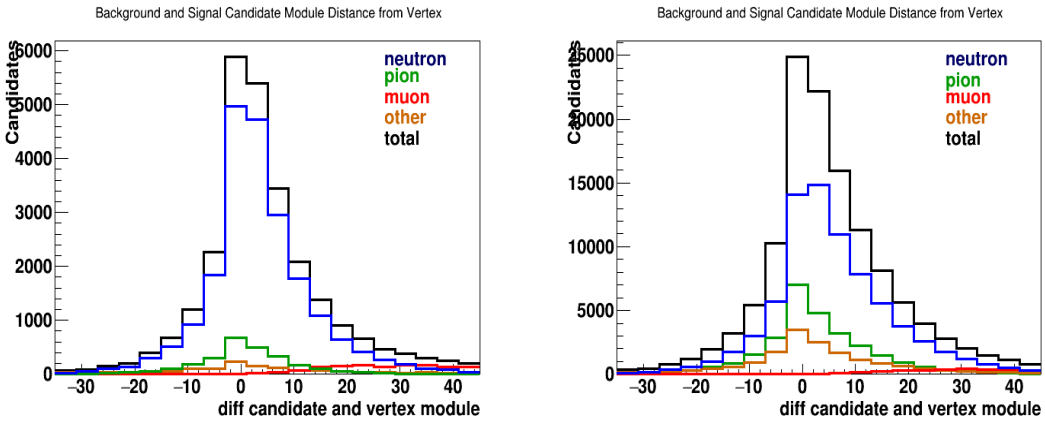

Figure 4.8: Plot of the difference between the candidate module and the interaction vertex module for the low q3 (left) and mid q3 (right) by what type of GENIE particle caused each candidate. Data points with corresponding statistical error bars are plotted. GENIE particle type includes neutrons(blue), negatively charged pions (green), muon fuzz (red), and other (orange). Muons cause almost all of the background candidates forward of the vertex.

these interactions. It was discussed previously that the background to the neutron candidates includes pions, protons, and photons coming from the muon. Figure 4.8 shows the transverse distance plots for the low q3 and mid q3 for each particle type that causes neutron candidates. The forward direction is primarily populated by signal neutrons. The backward direction is more a mix of signal and background. The muon background (the red line) populates a very specific area in the far forward direction. This would make sense since the muon itself only moves in the forward direction and thus any candidates produced from its interactions should be found here as well.

4.6 Full Transverse Position Distribution for Simulation in Comparison to Data

The distance a candidate is found in the transverse direction could indicate the origin of discrepancies between data and simulation. We calculated the distance in millimeters between the between the candidate and the interaction vertex strip in whichever view is the most appropriate for the neutron candidate, which may have activity in as little as one X, U, or V view. The neutrino interaction vertex position is a 3D object and has a well defined position in each of the three views. The transverse position distribution is made for three sub-samples: far upstream, far downstream, and near the vertex. Upstream or downstream candidates were 10 modules or more away from the interaction vertex in the corresponding direction. Candidates near the vertex were less then 10 modules away from the interaction vertex. These are shown in Figure 4.9.

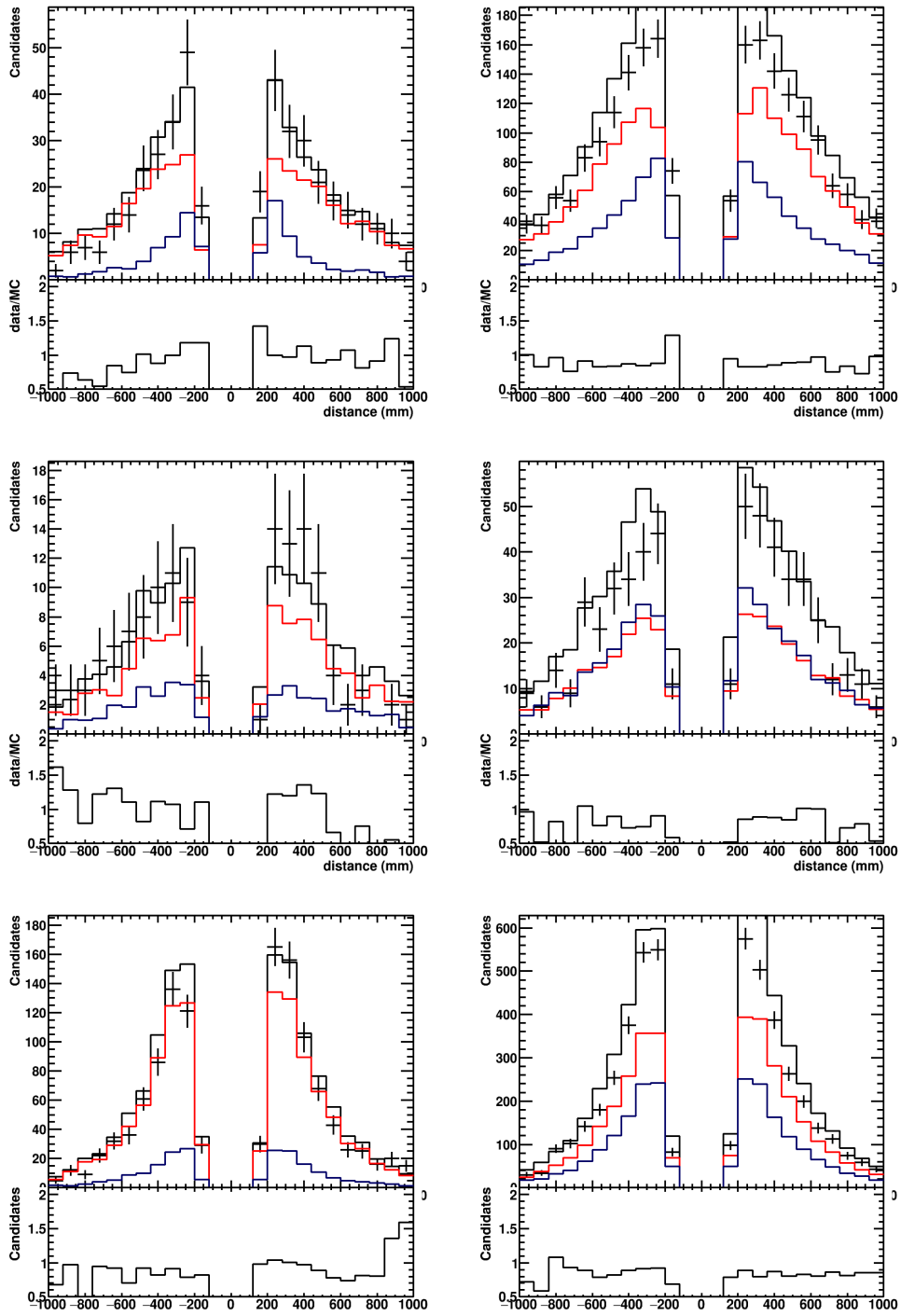

Figure 4.9: Plot of the transverse distance (mm) between the candidate and the interaction vertex for low q3 (left) and mid q3 (right) MC events separated by interaction type. The simulation includes RPA and 2p2h and the ratio compares the data and simulation to the default GENIE simulation. Data points with corresponding statistical error bars are plotted over top. Black is total, red is signal, and blue is background. Candidates are more then 10 modules upstream (top), more then 10 modules downstream (middle), or within 10 modules (bototm) of the interaction vertex.

The over simulation of candidates is stronger for the mid q3 than the low q3. This is true for candidates upstream, downstream, and near the vertex. It seems as though most of the disagreement in the low q3 is seen in candidates that move downstream of the vertex. The mid q3 is a different story with candidates in almost any bin of the three plots not agreeing with data, most often being over simulated.

Chapter 5

Comparison of Models to Data

Each section in this chapter is dedicated to a different model effect and discusses what the effect is meant to accomplish and what changes should be noticeable in the simulation. The distributions built in chapters 1, 3, and 4 are included here for each effect and are compared to what is found in the data. In each case, the question of interest is whether the model change helps the simulation describe the data, and/or reduces sensitivity to the RPA and 2p2h models.

For anti-neutrino interactions with carbon nuclei there are many discrepancies between the default simulation and the data taken with the MINERvA detector. These discrepancies are most easily seen in Figure 1.2, (repeated in the left column of Figure 5.1) which shows which shows the reconstructed available energy for both the low and mid momentum transfers. The simulation appears to under predict quite significantly in the hadronic energy region in between the quasi-elastic and delta interaction peaks in both the low q3 and mid q3 plots. There also appears to be an over (under) simulation near the quasi-elastic peak for the low (mid) q3. The analysis is aimed to figure out why the discrepancies appear by determining what types of interaction processes we are not simulating well and what is present in the data. The simulated interactions seen in the figure are adjusted during numerous systematic studies done using different weights simulated to account for known but uncertain effects we expect.

The plot set up seen in 5.1 is used throughout each of the sections in this chapter and the next. Here the left plots show the hadronic energy distributions, the middle plots show the neutron counting, and the right plots show the z-distribution of the neutron candidates. The top plots are all for the low q3 and the bottom plots are all for the mid q3.The ratio for the hadronic energy plots is always to the current simulation. The neutron counting and the z-distribution plots have a ratio taken with respect to the default simulation when only RPA and 2p2h are being tested. When a systematic is being tested, the ratio is to the model with RPA and 2p2h, so the deviation from this new baseline represents a component of an error band.

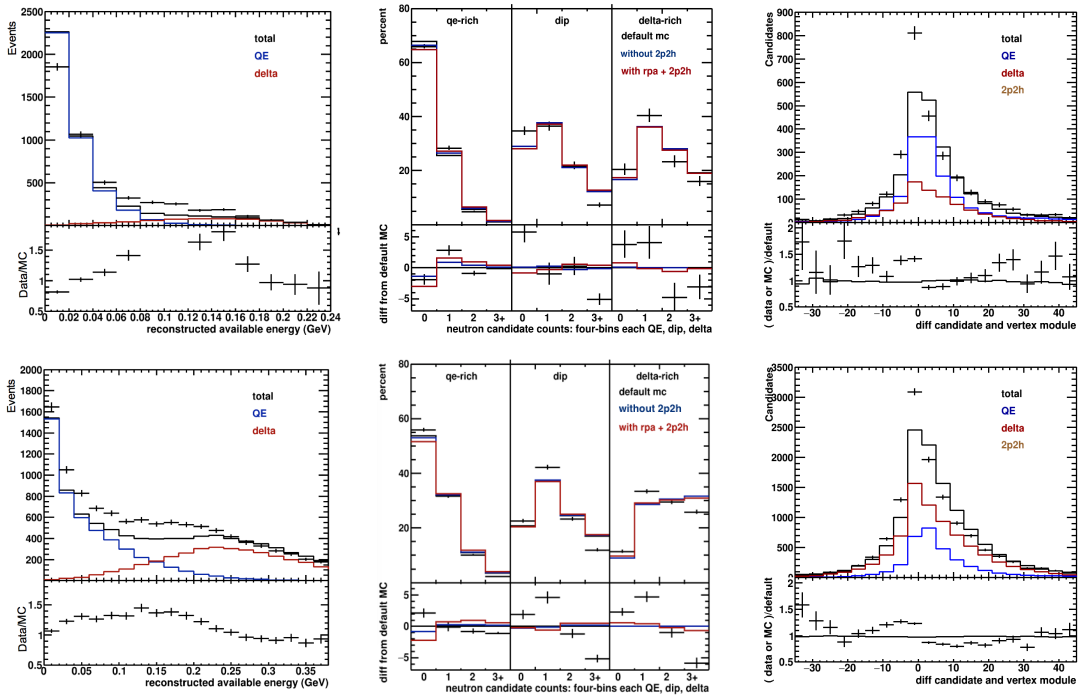

Figure 5.1: Default simulation. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and RPA plus 2p2h. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the default simulaiton while the ratio for the left plots is to the current simulation.

Different regions of these three distributions are affected by each model or systematic and each distribution is affected differently. In some regions the model moves in the direction of the data, but doesn't improve the agreement in other regions. The same can be said for the different distributions. This is because the different distributions offer different kinds of rate information. The middle plot contains only relative (percent) information, though because the data in the zero bin is higher for five of the six sub-panels, the simulations are overproducing neutron candidates. The left plot indicates if the overall event rate is off, while the right plot indicates discrepancies attributable to both the event rate and neutron production.

One example can been seen with the addition of the 2p2h events. The addition of these events would increase the simulation in the dip region for the hadronic energy plots bring simulation towards the data for both the low q3 and the mid q3. However for the neutron counting there is already an over simulation of events with neutron candidates in the dip region for both the low q3 and mid q3 so the addition of events with neutrons forces simulation further away from data. It is obvious the default simulation, the simulation with RPA, and then with RPA and 2p2h do not match the data in most bins in the neutron counting plots.

5.1 RPA Effect

The most distinct difference appears with the suppression of the first bin of reconstructed available energy with the addition of the RPA screening, significantly helping agreement between data and simulation in this area for the low q3. The opposite can be said for the mid q3 as this additional weight pulls the simulation further away from the data with the suppression of this first bin. This suppression is much larger in the low q3 than it is for the mid q3 and does not appear elsewhere in the reconstructed energy distributions.

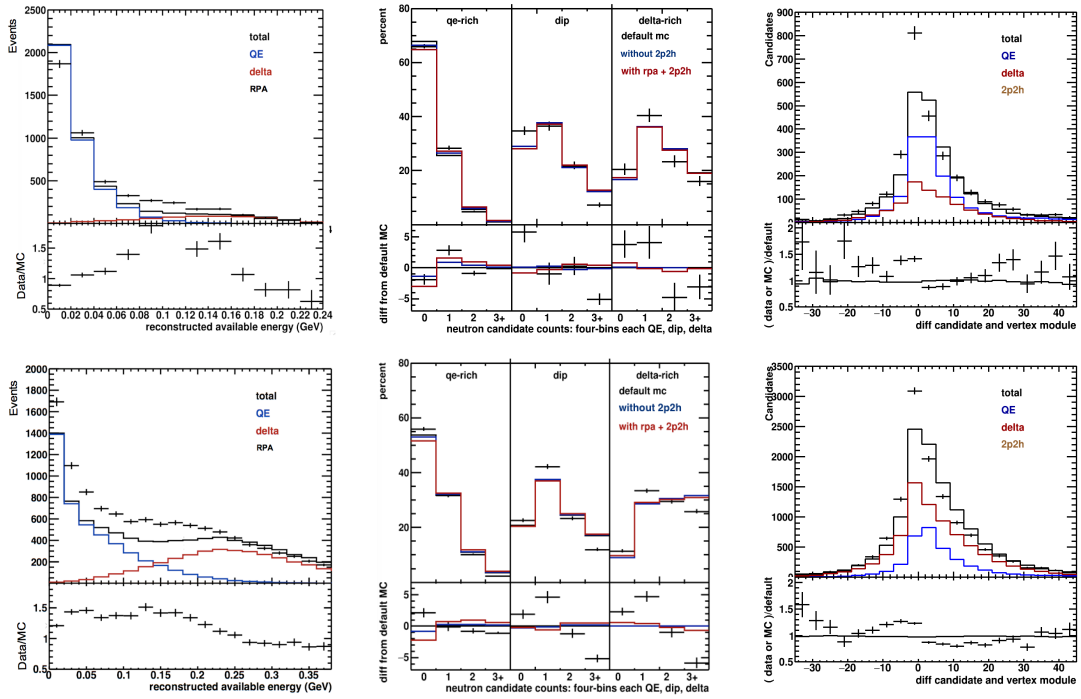

Figure 5.2: Simulation with RPA. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the default simulation while the ratio for the left plots is to the current simulation.

The addition of the RPA suppressing QE events within the first bin of the hadronic energy plots suggests there would be a decrease in events with the lowest energy neutrons or events with no visible neutron. This in turn suggests a decrease in the zero candidate bin for the QE region of the neutron counting plots in Figure 5.2 which is exactly what is found. Since the effect is greater in the low q3 hadronic energy plot than the mid q3 plot it is larger in the neutron counting plots for the low q3 as well. The addition of the RPA effect causes the zero candidate QE region bin to be suppressed by approximately 3 percent compared to only 1 percent change in the mid q3. The other bins in the QE region increase correspondingly, so each

sub-panel adds up to 100 percent. This causes a significantly better agreement in the QE region for the low q3 and just a slightly worse agreement for the mid q3. No noticeable change is seen in the bins of the dip or delta regions as the RPA only affects the quasi-elastic type interaction and very few of these spill into the dip and delta regions.

The RPA effect only causes a suppression in the QE event rate. However this suppression targets the lowest hadronic energy events with the least-visible neutrons, causing no significant change in the number of neutron candidates per event or their z-distributions.

5.2 The Valencia 2p2h Model

The 2p2h model adds events with at least two nucleons in the final state, often two neutrons. Before the addition of this process, Delta production followed by pion absorption was the main process with a similar two-nucleon final state. It is noticeable in the hadronic energy plots when the 2p2h process is added. These hadronic energy plots for the simulation with RPA and 2p2h for the low q3 and mid q3 are shown in 5.3.

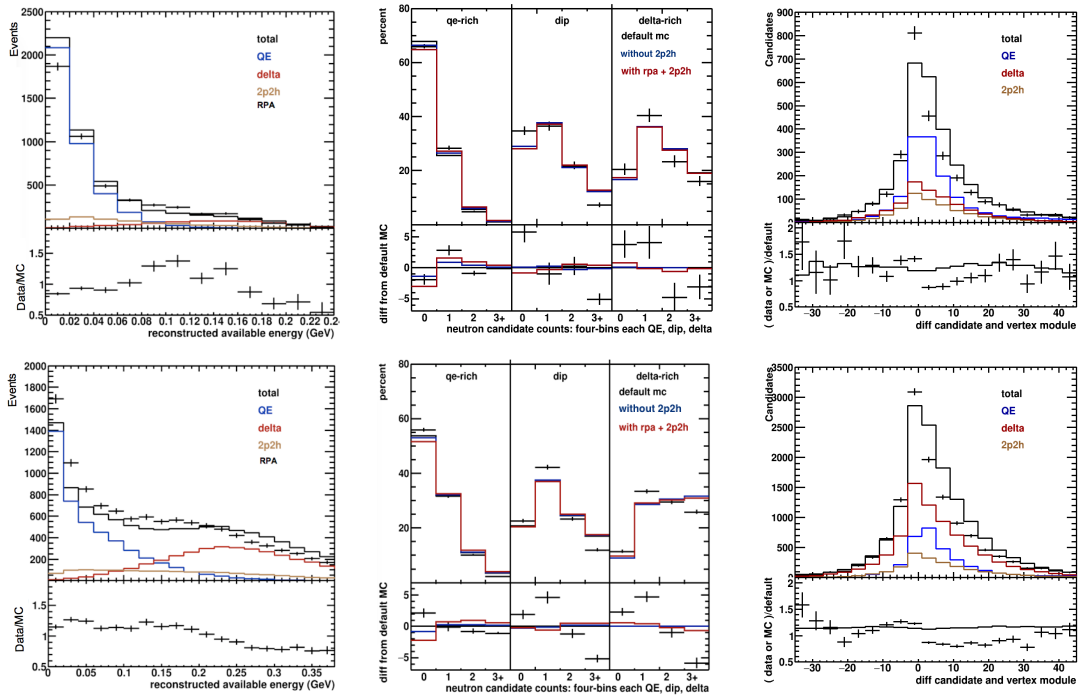

Figure 5.3: Simulation with RPA and 2p2h. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the default simulation while the ratio for the left plots is to the current simulation.

This addition causes the simulation in the region of hadronic energy in between the QE and delta events to increase, causing data and simulation to better agree. The addition of the 2p2h events increases the number of events with higher hadronic energy than seen for QE events so their neutrons will be more visible on average.

This increase of events with potentially visible neutrons inherently causes an increase in the events with two or more neutron candidates in them, and decrease in the percentage of events with zero candidates, possibly even one candidate. The trend stands out in the QE-rich region, comparing the red and the blue lines in Figure 5.3. The same trend, but smaller in magnitude is visible in the dip region. By itself, this effect does not improve the agreement with the simulation. Before the addition of the 2p2h it is obvious in the neutron counting plots in Figure 5.3 that the simulation already had more events with multiple neutrons in them then the data. The size of the change in the percentage of events with no candidates in the QE hadronic energy region being so large could be due to the fact the neutrons found here are rarely visible and the addition of events with two neutrons in them causes a significant increase in the number per event.

It is also interesting to note how small the changes are overall with the addition of the 2p2h. The addition of the 2p2h causes about a three percent change in the first bin in the QE region of the low q3 with the rest of the bins being a change of two percent or less, and similar for the mid q3 as well. For the mid q3 the overall neutron counting is very similar to that of the dip region itself because the mix of the QE and delta regions look similar to the dip region. Thus adding the 2p2h events, which fall primarily in this dip region and would have similar amounts of candidates per event as the total mid q3, gives an even smaller change than that in the low q3. If one looks at the low q3 hadronic energy plots it is apparent the QE region dominates meaning the overall count of neutrons in the low q3 should look very similar to the QE region alone. Since the addition of events is small in the low q3 the overall change in the low q3 is small but it is larger then what is seen in the mid q3. Tables 5.1 to 5.3 can aid with this discussion.

There is a slight increase in the zero candidate bins for the Delta region in both the low q3 and mid q3 which seems slightly suspicious. The guess is before the addition of the 2p2h events the delta region consisted of only delta events where a charged pion is likely to have come out of the nucleus and interacted in the detector along with a neutron that is more energetic then those found in the QE and dip hadronic energy regions. Charged pions make up a large portion of the background candidates counted by the algorithm and contribute to the total candidates here. The energetic neutrons found here are also the most likely in the entire sample to produce a candidate. The majority of the 2p2h events appear in the dip region when added into the simulation but the 2p2h events with a neutron and proton in the final state are more likely to fall in the delta region.

Low q3

Neutron Candidates	0	1	2	3+	% that is background
QE Region Events	67.83% ±0.25%	25.46% ±0.23%	05.63% ±0.12%	1.08% ±0.05%	16.03%
Dip Region Events	28.85% ±0.69%	37.41% ±0.74%	21.45% ±0.63%	12.29% ±0.50%	33.06%
Delta Region Events	16.61% ±0.67%	36.23% ±0.87%	28.09% ±0.81%	19.07% ±0.71%	44.19%
All Low q3	60.28%	27.42%	8.82%	3.48%	

Mid q3

Neutron Candidates	0	1	2	3+	% that is background
QE Region Events	53.81% ±0.27%	31.84% ±0.25%	10.92% ±0.17%	3.43% ±0.10%	13.92%
Dip Region Events	20.71% ±0.27%	37.60% ±0.32%	24.54% ±0.29%	17.14% ±0.25%	35.05%
Delta Region Events	09.14% ±0.16%	28.70% ±0.26%	30.47% ±0.26%	31.70% ±0.26%	58.95%
All Mid q3	29.83%	32.20%	21.20%	16.77%	

Table 5.1: Percentage of events that had 1,2,or 3+ candidates in them along with the statistical uncertainty for each region of hadronic energy: QE, Delta, and Dip. Made for the default simulation.

Low q3

Neutron Candidates	0	1	2	3+	% that is background
QE Region Events	66.44% ±0.26%	26.39% ±0.24%	5.99% ±0.13%	1.18% ±0.06%	15.63%
Dip Region Events	28.95% ±0.69%	37.69% ±0.74%	21.21% ±0.62%	12.15% ±0.50%	32.69%
Delta Region Events	16.67% ±0.67%	36.22% ±0.86%	28.06% ±0.81%	19.04% ±0.71%	44.25%
All Low q3	58.63%	28.35%	9.30%	3.72%	

Mid q3

Neutron Candidates	0	1	2	3+	% that is background
QE Region Events	53.04% ±0.28%	32.13% ±0.26%	11.22% ±0.18%	3.61% ±0.10%	13.88%
Dip Region Events	20.44% ±0.27%	37.54% ±0.33%	24.69% ±0.29%	17.32% ±0.25%	35.35%
Delta Region Events	9.10% ±0.16%	28.68% ±0.26%	30.48% ±0.26%	31.74% ±0.26%	58.90%
All Mid q3	28.52%	32.28%	21.76%	14.45%	

Table 5.2: Percentage of events that had 1,2,or 3+ candidates in them along with the statistical uncertainty for each region of hadronic energy: QE, Delta, and Dip. Made for the simulation with RPA.

Low q3

Neutron Candidates	0	1	2	3+	% that is background
QE Region Events	64.82% ±0.25%	27.12% ±0.23%	6.61% ±0.13%	1.45% ±0.06%	14.47%
Dip Region Events	28.14% ±0.57%	37.09% ±0.62%	22.07% ±0.53%	12.69% ±0.42%	24.57%
Delta Region Events	17.37% ±0.25%	36.05% ±0.29%	27.53% ±0.73%	19.04% ±0.64%	46.94%
All Low q3	56.15%	29.15%	10.34%	4.36%	

Mid q3

Neutron Candidates	0	1	2	3+	% that is background
QE Region Events	51.59% ±0.27%	32.51% ±0.25%	11.90% ±0.17%	3.99% ±0.10%	12.90%
Dip Region Events	20.45% ±0.27%	36.98% ±0.29%	25.02% ±0.26%	17.55% ±0.23%	30.16%
Delta Region Events	9.66% ±0.16%	29.08% ±0.24%	30.27% ±0.24%	30.99% ±0.25%	55.10%
All Mid q3	27.81%	32.49%	22.17%	17.53%	

Table 5.3: Percentage of events that had 1,2,or 3+ candidates in them along with the statistical uncertainty for each region of hadronic energy: QE, Delta, and Dip. Made for the simulation with RPA and 2p2h

The percentage of the candidates in simulation that are background, which is determined by the truth information, and the statistical uncertainties for each of the candidate count bins can be found in Tables 5.1 to 5.3. The same information seen in the neutron counting plots along with the percentage of events in each region, as well as the whole low q3 or mid q3, for the default simulation, with RPA, with RPA and 2p2h each by themselves is also shown in Tables 5.1 to 5.3. The tables are highlighted to show where the change in each bin from the previous table is significant enough to be outside the statistical error bars. These bins show where the simulation is sensitive to adding in the 2p2h. Of particular interest, the simulation is most sensitive to adding in the 2p2h in bins of neutron candidates per event that are already over simulating neutrons pushing the simulation even further from the data.

The information shown in Tables 5.4 and 5.5 gives the total number of events there are in the data and the simulation, after being normalized to the data, with either 0, 1, 2, or 3 plus candidates in them. These totals can be used with the previous tables to help understand the overall percentages and a magnitude of the number of events lost or gained due to the uncertainty in each of the bins. The total number of candidates in each hadronic energy region is provided as well. While the data may contain slightly more events in some regions there are many more data events causing the total number of candidates per event for the simulation to be significantly greater.

Candidates	0	1	2	3+	Total Events	Total Candidates
QE-rich Data	2364	1010	180	36	3590	1478
QE-rich MC	2587	1082	263	60	3012	1788
Dip Data	297	322	177	66	862	874
Dip MC	189	249	149	85	672	802
Delta-rich Data	93	162	100	59	414	539
Delta-rich MC	67	137	106	73	383	568

Table 5.4: Comparison of total events in data and simulation with 0, 1, 2, or 3+ candidates for the low q3 with RPA and 2p2h turned on.

.

Candidates	0	1	2	3+	Total Events	Total Candidates
QE-rich Data	2524	1443	460	107	4534	2684
QE-rich MC	1985	1250	458	153	3846	2625
Dip Data	745	1405	778	399	3327	4158
Dip MC	599	1084	734	514	2931	4094
Delta-rich Data	361	1069	958	816	3204	5433
Delta-rich MC	371	1117	1162	1190	3840	7011

Table 5.5: Comparison of total events in data and simulation with 0, 1, 2, or 3+ candidates for the mid q3 with RPA and 2p2h turned on.

5.3 Hadronic Energy Scale Uncertainty

There is a limit in the sensitivity to the RPA and 2p2h effects caused by other systematic uncertainties present in the model. In the anti-neutrino case, the addition of the RPA and 2p2h does not describe the data, so one or more systematic effects must also be important.

Here the neutron counting plots are exactly the same as those seen in the 2p2h and RPA sections of this chapter except the differences shown in the bottom of the plots are now to the simulation with RPA and 2p2h instead of the default. The best way to understand this is to imagine a line at zero representing what the simulation with RPA and 2p2h looks like in comparison to any of the other four sets of information plotted. These plots will continue with the same scheme throughout the paper for any systematic as an addition to RPA and 2p2h, so it appears the way an error-band would. The lines without 2p2h and the default (without RPA or 2p2h) allow the reader to judge the error band magnitude relative to these changes in the model.

The MINERvA collaboration ran a miniature replica of the detector in a test beam to measure the energy response to protons, pions, and electrons. The uncertainty in the proton and pion hadronic energy, including both the Geant4 simulation and the detector response, was found to be about four percent meaning the hadronic energy plotted in previous sections could be off by an increase or decrease of that four percent. The test beam setup did not allow for a similar test with neutrons. However, since the number is known for pions and protons the hadronic energy calculated for events in this analysis can at least be increased or decreased by four percent. Also, since neutrons are expected to have an even larger uncertainty then the protons and pions, and constitute a large fraction of the final state in these events, we also tested a ten percent uncertainty.

This change explicitly moves events from left to right in the reconstructed energy plots, lowering the left-most peak in the simulation. This change also causes events to migrate between q3 slices because the change is made before the momentum transfer is calculated. The first plots seen in Figure 5.4 are for the hadronic energy increase of ten percent. This increase should shift events from lower slices of q3 to higher slices of q3 which means migrating events from the low q3 to the mid q3 and from the mid q3 to higher slices that are not analyzed here.

This increase should also cause events with hadronic energy sitting near the edges between hadronic energy regions (QE-rich, dip, Delta-rich) to shift from the lower one to the higher. These events with hadronic energy sitting on the high end of each hadronic energy region should be the events with the most visible neutrons in them because neutron visibility is most strongly correlated with neutron energy. This means the events with neutrons should be the ones leaving one hadronic energy region for the higher one. However, the events that sit in the zero candidate per event bins in the neutron counting plots shouldn't have much hadronic energy in them and would rarely migrate from one slice of q3 or hadronic energy region to another. Simply put the events without neutrons should stay and events with neutrons will migrate to higher energy and q3 regions. However the amount of events leaving and entering in each region are not symmetric causing a noticeable change.

Increasing the hadronic energy by ten percent causes the hadronic energy plots in Figure 5.4 for both the low q3 and mid q3 to have fewer events overall in comparison to the plots for the simulation with just RPA and 2p2h in Figure 5.3. This decrease causes the QE region in the low q3 and the delta region in the mid q3 to move towards data. However the decrease pulls simulation from the data in the QE region for the mid q3 as there was already an under-simulation of events existing here.

In the neutron counting plots there is an increase in the percentage of events without neutrons in them for both the low q3 and mid q3 as well as all three of the hadronic energy regions. This causes a significant improvement in matching the neutron counts between data and simulation for the low q3. Many of the bins in each of the three regions either agree with or are close to what the data shows. This shift is larger than the effect of adding RPA and 2p2h in most panels, and yet data suggest an even larger systematic shift could be appropriate.

The z-distance candidate distribution plots and the hadronic energy plots shown in Figure 5.4 appear to have fewer candidates than in 5.3, as they should since the

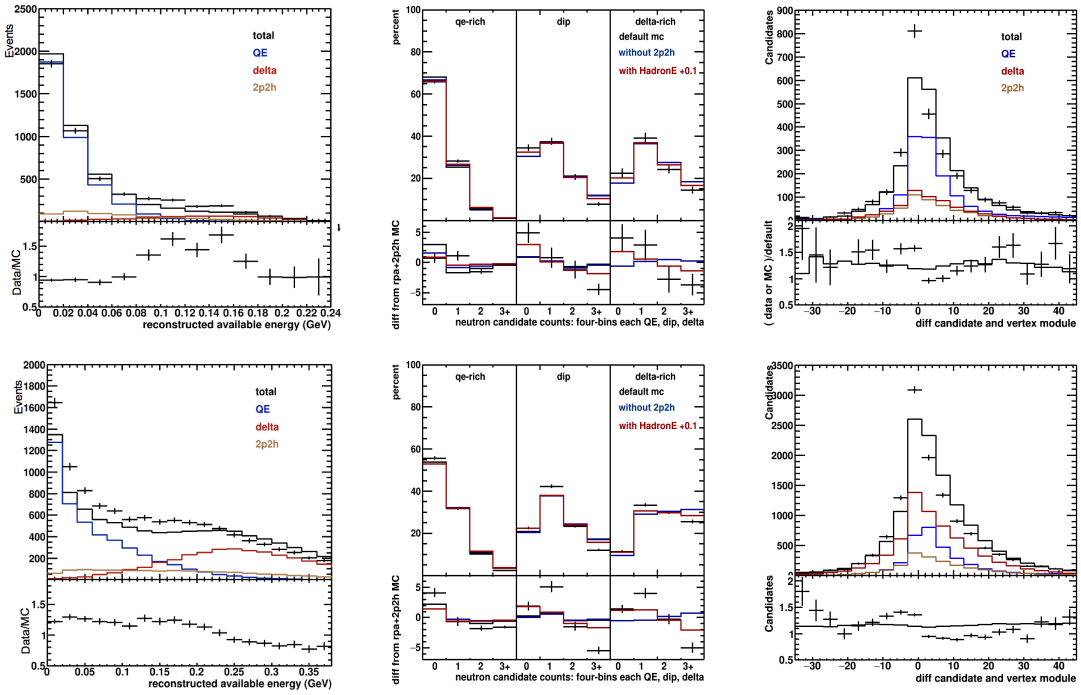

Figure 5.4: Simulation with RPA, 2p2h, and hadronic energy 10% increase. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the simulation with RPA and 2p2h while the ratio for the left plots is to the current simulation.

events are moving to higher slices of q3. The changes are small. More physically, the two distributions disagreed in the same way, so moving events from the low q3 to the mid q3 doesn't improve the agreement. However, the implementation of this hadronic energy scale systematic might not be true to the origin of a systematic mistake, if it is to do with the number of neutrons or the probability that they leave energy in the detector. Testing or fixing something more fundamental could produce a change not seen in this test.

The story for the decrease in the hadronic energy by ten percent is similar as to that of the increase of ten percent. Figure 5.5 includes all of the plots for the decrease in the hadronic energy by ten percent. This decrease should pull events from higher slices of q3 and hadronic energy regions into the lower q3 slices and hadronic energy regions. These events should once again have neutrons in them because those with out neutrons should have little energy and a small change to something small would not give a significant change. However now we are adding the events with neutrons to the regions studied in this analysis, instead of taking from them as with the increase of ten percent, causing an increase in the percentage of events with neutrons in them. This change is once again small but is now in the

opposite direction of the ten percent increase and thus causing more disagreement between the data and the simulation when it comes to neutron counting. The same can be said for the z-distance distribution and hadronic energy plots.

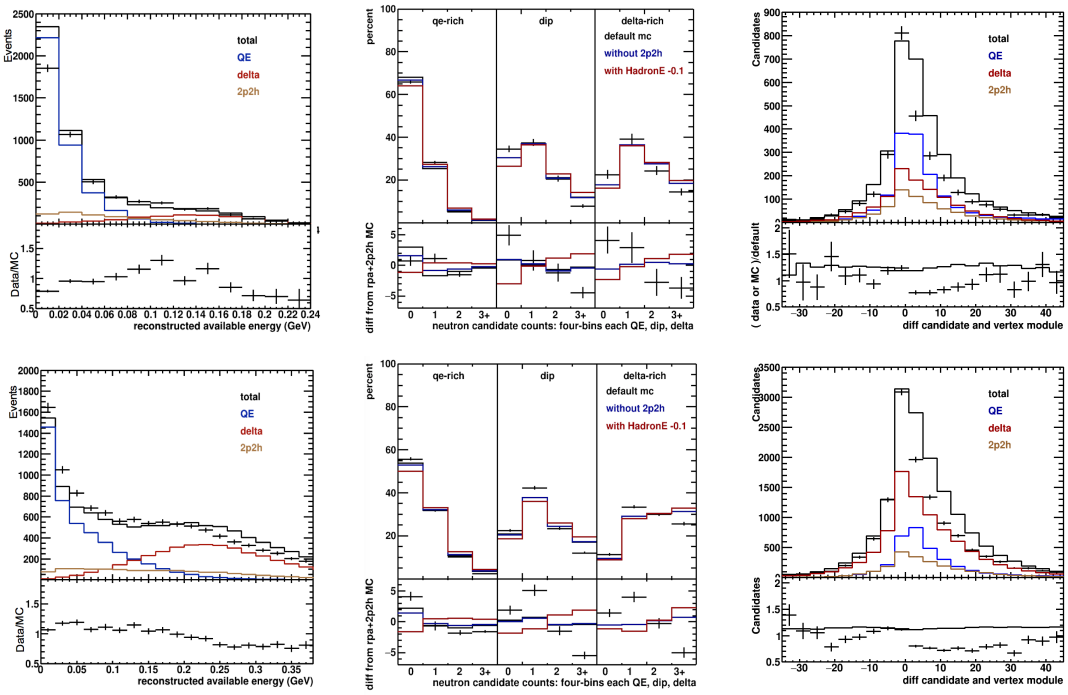

Figure 5.5: Simulation with RPA, 2p2h, and hadronic energy 10% decrease. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the simulation with RPA and 2p2h while the ratio for the left plots is to the current simulation.

Since the increase of ten percent does a good job in matching neutron counts between data and simulation, and for the sake of brevity, the four percent increase and decrease plots have be omitted. The magnitude of the changes in each of the three distributions for both the low q3 and mid q3 are about 1.5 times smaller then those seen for the ten percent increase or decrease. These changes also follow suit by moving in the same direction as those for the ten percent changes.

5.4 Increase Pion Absorption for Delta Production

This weight doubles delta events where the pion is absorbed into the nucleus knocking out a second nucleons. It's a standard process that is very similar to the 2p2h events and it could even be considered a limiting case of a 2p2h event. In turn

it also decreases the events where this does not occur meaning a pion along with one nucleon leave the nucleus. This increase and decrease of different event types was normalized in a way such that the overall event rate did not change. The origin of this test, John Demgen's thesis, was to enhance two-nucleon final states in the dip region using something other than the 2p2h process.

The hadronic energy plots in Figure 5.6 show an increase that is mainly noticeable within the dip region of the mid q3. This is due to the type of events being weighted up. The delta events with two nucleons coming from the nucleus appear very much like 2p2h events would and have similar hadronic energy thus they appear in the same spot in the hadronic energy plots. This inclusion causes the bins in the dip region of the mid q3 to move closer to what is seen in the data.

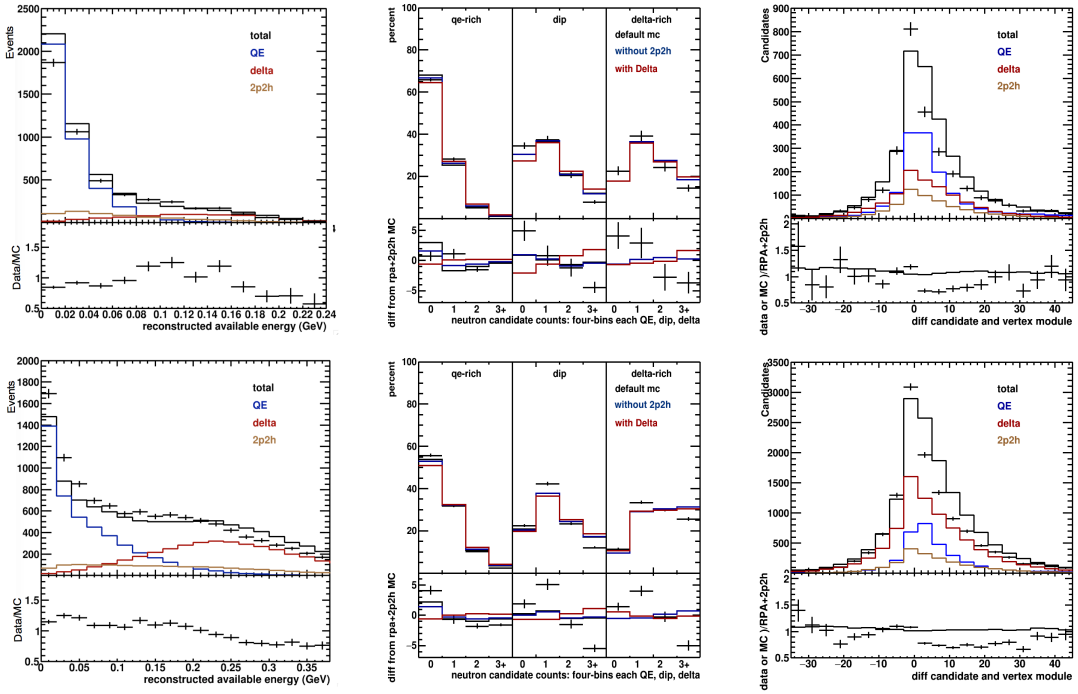

Figure 5.6: Simulation with RPA, 2p2h, and Delta. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the simulation with RPA and 2p2h while the ratio for the left plots is to the current simulation.

For neutron candidate counting this change in the delta events causes less agreement between data and simulation. Due to the addition of events with two nucleons, neutrons and protons potentially able to interact in the detector and create candidates, it would be expected for the percentage of events with visible candidates to increase. For either slice of momentum transfer there is a decrease in the zero candidate bin for both the QE and dip regions followed by an increase in 2 or 3 plus

candidates bins. The delta appears to follow an opposite trend which is hard to describe but could have a similar solution as to the one discussed for the 2p2h addition earlier on. This however is forcing the simulation even further from the data because the current simulation seems to over produce the number of candidates there should be per event, when comparing to data, and thus only continues to add additional candidates. Reducing pion absorption after Delta production could better describe the data, and there may be some double counting of processes with the 2p2h model we have added.

Increasing the pion absorption causes the two-nucleon final states that are likely to produce one or two neutron candidates to be more prominent then before. This causes an increase in the overall total number of neutron candidates which is obvious from the z-distribution plots. Similar to the other systematics, the candidates are not placed in any particular part of the z-distribution plot and only forces simulation further away from the data.

5.5 Increase the QE Axial Mass Parameter

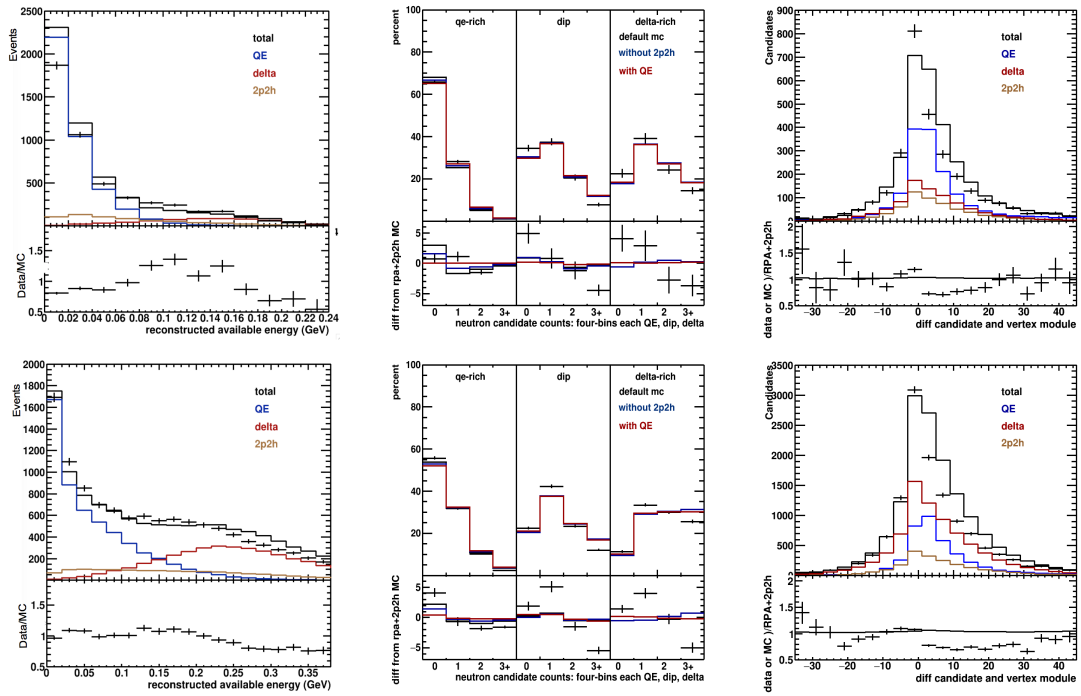

Figure 5.7: Simulation with RPA, 2p2h, and QE. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the simulation with RPA and 2p2h while the ratio for the left plots is to the current simulation.

The QE axial mass is a parameter extracted from neutrino-deuterium measurements [5]. Increasing the parameter increases the number of QE type interaction events; the effect is larger for higher three-momentum transfer. The uncertainty being applied here is the default GENIE uncertainty of +25%. Given the discrepancy seen in the hadronic energy for the mid q3 and low q3 plots in Figure 1.7 it was beneficial to see how the addition of it would affect the neutron counting and the hadronic energy plots. Figure 5.7 includes the hadronic energy, neutron counting, and z-distance modules plots for both the low q3 (top) and mid q3 (bottom) with the new QE weight being turned on.

The first bins in the hadronic energy plots for both three-momentum transfer slices change quite dramatically from what the default simulation show with the addition of the QE weight. This increase of the QE interaction rate helps the simulation describe the data for the mid q3 as the simulation was much lower then the data, yet the opposite is true for the low q3. Here there was already a surplus of QE events and this addition only increases that difference.

The addition of the QE weight to the RPA and 2p2h simulation has little effect on neutron counting. In the neutron candidate counting plots of Figure 5.7 the data and the simulation with the QE weight plus RPA and 2p2h are represented. The hadronic energy plots naively suggest the neutron counts for the QE region in the low and mid q3 would change significantly. However the increase in QE events is proportional throughout the entire region, adding in proportional numbers of events with and without visible neutrons to the previous total, meaning the percentage of events with and without candidates does not change.

The increase in the QE events does cause an increase in the overall total number of candidates. This increase is present in the z-distribution plots shown above given the QE line has increase significantly. This increase is moving simulation in the opposite direction needed to help it describe the data given the over production of candidates from the model with RPA and 2p2h alone.

5.6 Flux Systematic

The flux weight used in this part of the analysis decreases anti-neutrinos with energies less then 4 GeV (this includes nearly the entire sample discussed in this paper) which distorts the flux shape and simply decreases the overall event rate. Due to this the neutron candidate counting plots are expected to not change with the addition of the weight. The hadronic energy plots should only show a decrease in overall events but the shape of the simulation and the ratio of data to simulation should stay the same. Similar to the other systematics discussed this flux change is an addition to simulation with the RPA and 2p2h turned on.

The plots shown in Figure 5.8 validate exactly what was expected. The total neutron candidates counted for each hadronic energy region do not change. The hadronic energy plots show a lower overall event rate but the overall shape has remained the same. This shape would only have changed had each event type been weighted dis-proportionally to another which is not what the flux weight does. The decrease in the number of simulated events also decreases the total number of

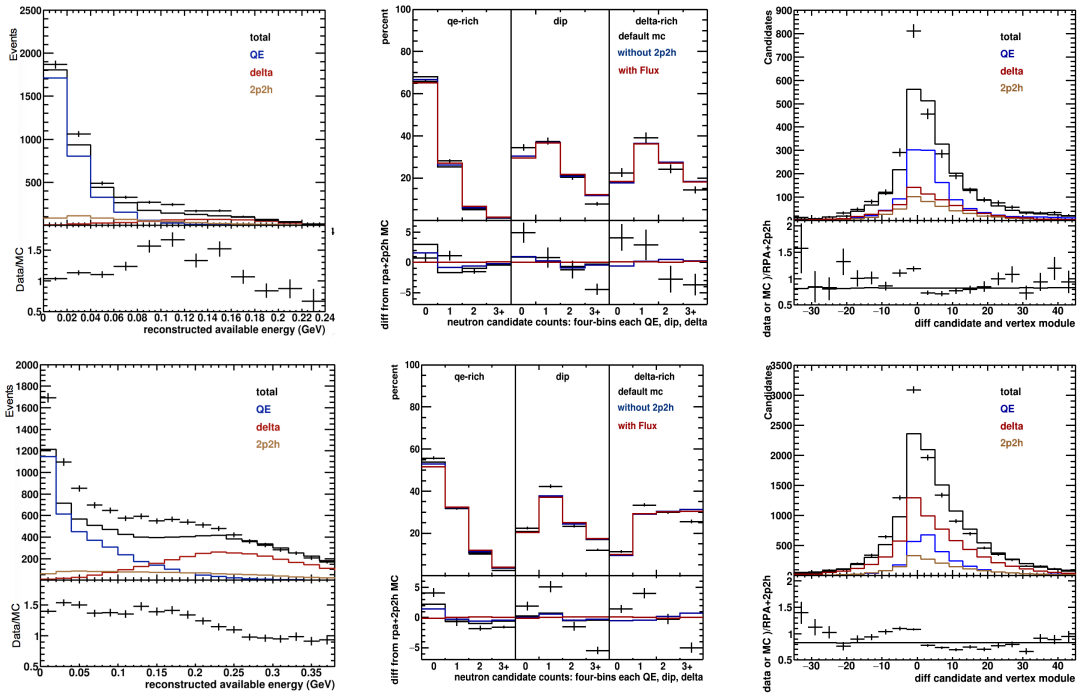

Figure 5.8: Simulation with RPA, 2p2h, and Flux. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the simulation with RPA and 2p2h while the ratio for the left plots is to the current simulation.

candidates produced but does not force the z-distribution plots for the data and the simulation to match. While it was anticipated it is good practice to check that what is expected actually happens. However this adjusted flux does not help compensate for the differences between the simulation and the data.

5.7 Intra-nuclear Re-scattering Mean Free Path

This weight changes the mean free path for nucleons and pions that re-scatter before leaving the nucleus causing the final states we see in the detector to change. The nucleon mean free path change will increase reactions and decrease events with no reaction, or vice versa. If all events with and without re-interactions are similar enough and thus weighted similarly a change won't be very noticeable. The increase of the nucleon mean free path is shown in Figure 5.9. Both the low q3 and mid q3 show very little change due to this.

This increase in the nucleon mean free path does not have a huge impact on the hadronic energy. The changes are noticeable while looking closely at the plots, but overall the plots remain the same. The increase in the nucleon intra-nuclear re-scattering mean free path should cause an increase in the number of neutron can-

didates. This is because the protons and neutrons interacting inside of the nucleus would knock out neutrons and potentially more protons that could produce neutron candidates.

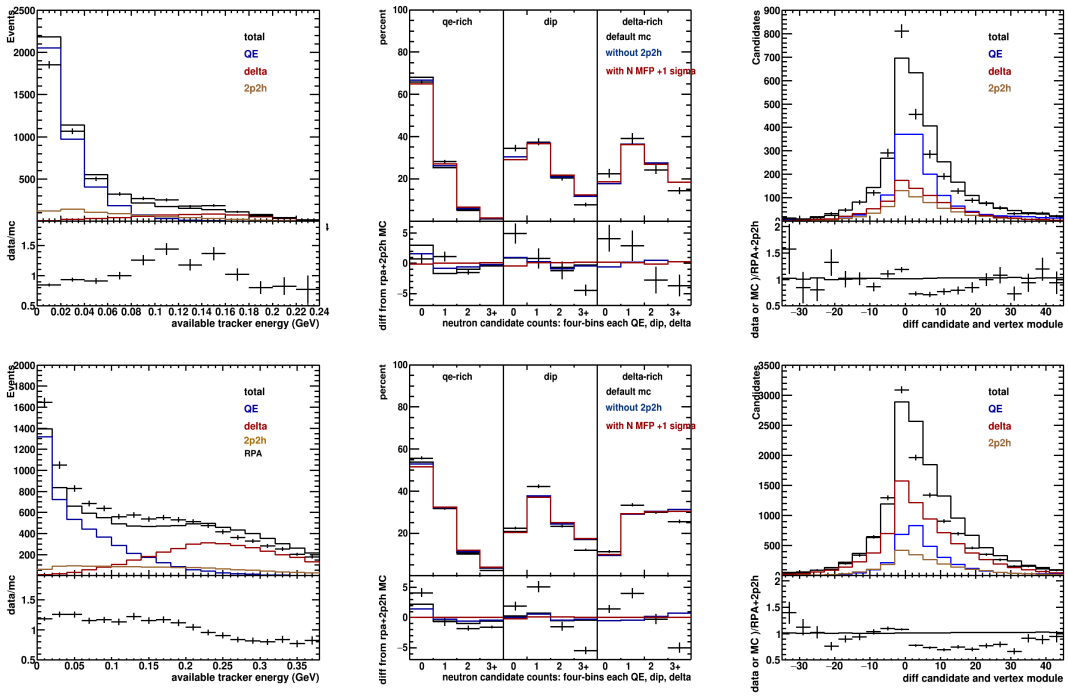

Figure 5.9: Simulation with RPA, 2p2h, and Nucleon Intra-Nuclear MFP Increase. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the simulation with RPA and 2p2h while the ratio for the left plots is to the current simulation.

The effect on the z-distribution is also very small. There is a small increase in the percentage of events with neutron candidates in them, this means there shouldn't be a large amount of candidates being added to the sample. Changes are only noticed if studied carefully between the z-distributions observed in Figure 5.9 and the z-distributions for the RPA and 2p2h only in Figure 5.3. In fact, this is the only systematic that seems to change the shape of the z-distributions (see the low q3 ratio) in ways that resemble the feature seen in the data.

Figure 5.10 includes the plots for the nucleon mean free path being decreased by the one sigma value assigned to it. The addition of this weight once again causes very little change in the plots seen here for similar reasons as discussed in the previous paragraphs.

There is a separate systematic for pion re-interactions. The pions in the anti-neutrino sample come from events with the delta interactions that appear mainly in the delta hadronic energy region with some spilling into the dip region. The pions in

these events cause between 10 and 20 percent of the total neutron candidates counted for either the low q3 or mid q3 respectively. These candidates are then spread out among the six bins in the dip and delta regions that contain candidates. These bins should contain most of the neutron candidates in the anti-neutrino sample since the dip and delta regions should also contain the most energetic and visible neutrons. If these candidates fall within the 6 bins with the most candidates, and get small weight adjustments, then the overall percentage of events with candidates should not have a noticeable change when the pion intra-nuclear MFP is adjusted up and down. The plots agree with this statement as seen in Figure 5.11 and Figure 5.12.

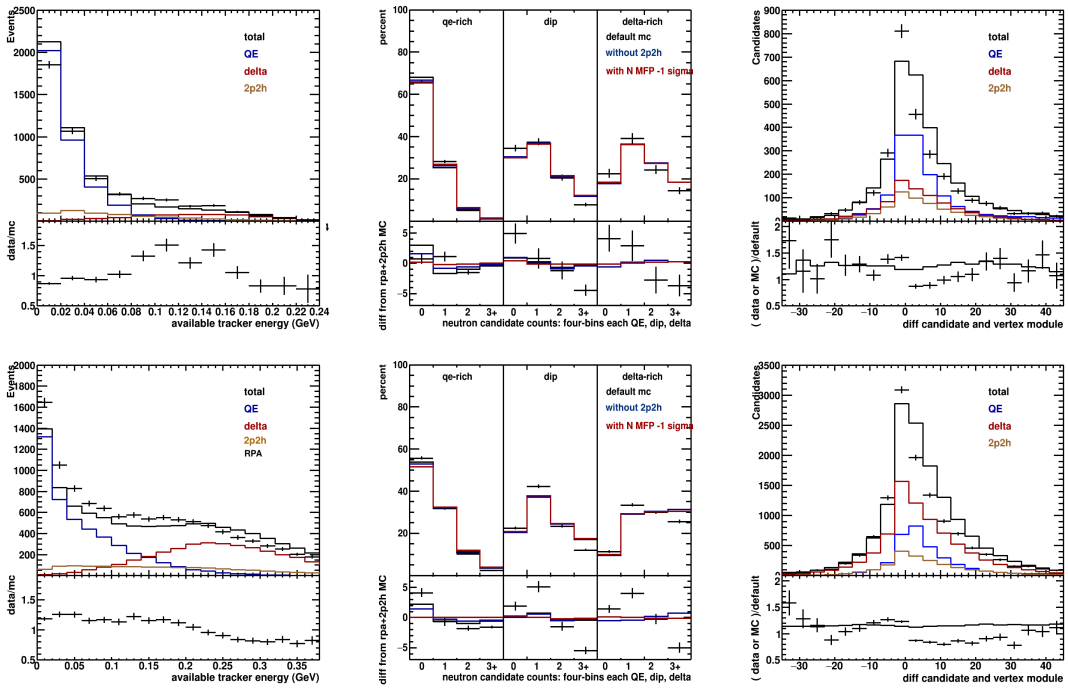

Figure 5.10: Simulation with RPA, 2p2h, and Nucleon Intra-Nuclear MFP Decrease. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the simulation with RPA and 2p2h while the ratio for the left plots is to the current simulation.

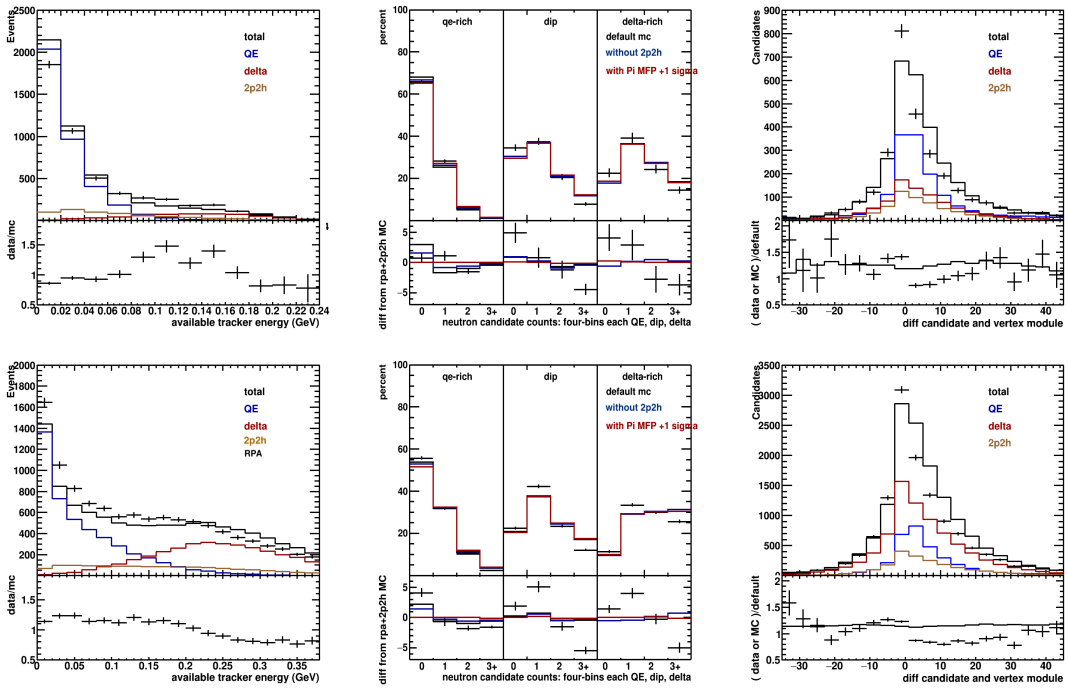

Figure 5.11: Simulation with RPA, 2p2h, and Pion Intra-Nuclear MFP Increase. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the simulation with RPA and 2p2h while the ratio for the left plots is to the current simulation.

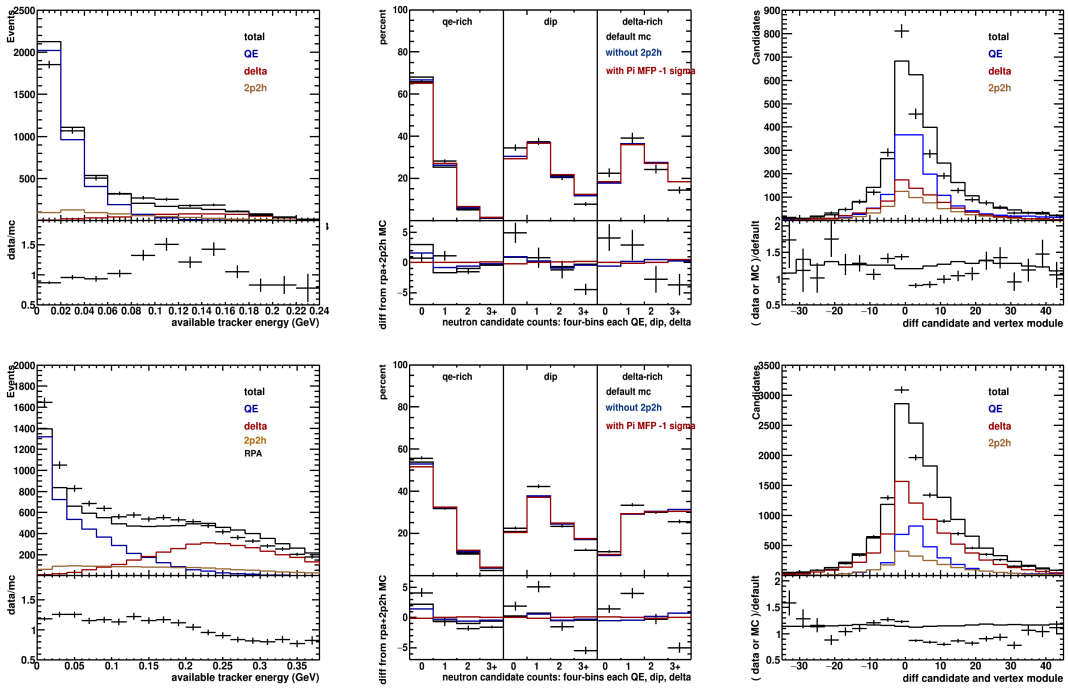

Figure 5.12: Simulation with RPA, 2p2h, and Pion Intra-Nuclear MFP Decrease. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the simulation with RPA and 2p2h while the ratio for the left plots is to the current simulation.

5.8 Hadron Cross Section Changes

The following study is still in progress. It has yet to be determined if the process of cross section changes here match the needs of this analysis.

Neutron counting was completed using the Hadron Re-weight Tool, created by Jeffery Kleyklamp and fully described in the appendix of this thesis. This weight is equivalent to changing the cross sections used by Geant4 when it steps a neutron (or proton or pion) through the detector deciding whether it deposits energy or not. The re-weight tool can increase and decrease the inelastic and elastic (also called non-interacted) cross sections of protons, neutrons, and pions by MINERvA's estimate of the one-standard deviation uncertainty. [10]. The cross sections are increased or decreased by weighting events depending on how far a particle traveled before interacting. Each of these effects were done separately and impose an additional weight to the simulation on top of the RPA and 2p2h effects which are included as well.

A setting that mimics an increase in the inelastic cross section is shown in Figure 5.13. These weights have the side effect that the more common elastic interactions

get weighted down. When all of the normalizations are accounted for (the total number of events must stay the same), the counter-intuitive net effect of a decrease in the number of neutron candidates and an increase in events in the zero-bins occurs. While the changes are similar between the low q3 and the mid q3 they are not between hadronic energy regions.

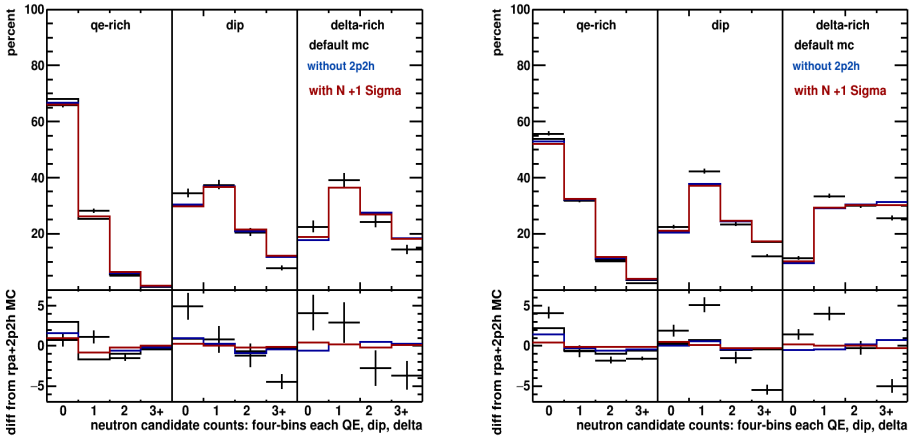

Figure 5.13: Percentage of neutron candidates per event while weighting the neutron cross section up for the low q3 (left) and mid q3 (right) for the QE, dip, and delta regions.

The QE hadronic energy region is almost completely all QE type events which always have a neutron in the final state. These neutrons are visible at most half of the time, mainly due to the quasi-elastic neutron re-interactions that are counted as non-interacted. If the events with visible and invisible neutrons interacting in them are weighted down due to being considered elastic, but differently because the path length for the different interactions will cause different weights, the effect here should be very noticeable and is. There is an increase in the percentage of events with no candidates in them with a similar decrease in events with one candidate. There are very few events with more then one candidate here because of the one neutron final state probability in the QE region so the corresponding bins are not affected by a large change.

The dip and delta regions have slight increases in the percentage of events with zero candidates in them. The magnitude of the change is smaller in these two sections in comparison to the QE region. The dip region has a smaller change likely due to the two neutron final state. Here it is probable the events with two neutrons in the final state would have one of them visible and possibly both. Since nearly all visible neutrons interact in the same way this causes almost all of the neutrons to be weighted down and in similar ways. After re-normalization this causes only a slight increase in the percentage of events with zero candidates. The delta region does not have two neutron final states but it does have energetic neutrons in the final state which are the most likely to be visible. If this is true for a majority of the events, and they are all quasi-elastic interactions with similar weights, then the same argument can be made here.

Adjusting the proton cross section does not change the neutron counting plots. Protons are a very small part of the background to the neutron counting for both the low q3 and the mid q3. Referencing Table 3.3 it can be seen that protons create about 0.28 percent of the neutron candidates in the low q3 and 1.54 percent in the mid q3. The proton cross section change has a max at 20 percent. So a 20 percent change in the 0.28 or 1.54 percent of events with the proton caused neutron candidates in them is negligible, especially because these events do not populate just one bin in the neutron counting plots.

The pion background is much larger than that of the proton background as pions make 10.79 percent of the neutron candidates in the low q3 and 22.31 percent in the mid q3. However, even though this background is much higher the effect of weighting the cross section of the pions either up or down is once again negligible. As before the maximum value the cross section of the pion can change is by 20 percent in each event. This, however, is unlikely as the 20 percent change is associated with the pion interacting in lead which would put the interaction outside of the tracker part of the MINERvA detector. Most of the pions interact within the carbon nuclei and thus the change in the cross section is approximately 10 percent. A 10 percent change in 10.79 percent of the candidates in the low q3 and even 22.31 percent of the candidates in the mid q3 still gives a small change. When the events with these pion caused candidates are spread out over multiple bins in the dip and delta regions the overall percentage of events with a certain number of candidates show minimal change.

Chapter 6

Proton Counting

Protons are a common product from anti-neutrino interactions, just less-so than neutrons. Fortunately protons in the detector are well calibrated and simulated, including tuned response and constraints from MINERvA's own test beam data. This observable offers additional complementary information to the three distributions described in the neutron counting analysis.

This chapter dives into the comparison between data and simulation while counting solely proton candidates in the anti-neutrino events in hopes of better identifying why the simulation and data do not agree. The same systematic studies as mentioned in chapter 4 were completed once again. An additional systematic on the Birks' law parameter, implemented only for protons, is included. The hadronic energy plots for these studies will change as they did in the neutron section (because the sample is identical), but are repeated next to the proton multiplicity distribution. Unlike neutrons, if there is a proton in the event, practically all its energy will be included in the hadronic energy plot.

6.1 Proton Counting Algorithm

The proton counting algorithm used for the following studies is unchanged from previous University of Minnesota - Duluth (UMD) student John Demgen, and was further developed by Alec Lovlein and Jake Leistico, and used for the publication [7]. As masters students at UMD, Ethan Miltenberger and John Demgen, worked on the sister analysis of this thesis which used neutrino interactions with carbon and hydrogen nuclei instead of anti-neutrino. From the conservation of charge the most prevalent products of these interactions were protons, instead of the neutrons found in the anti-neutrino case. The proton counting algorithm was created specifically to look for multi-proton final states. [12] [9].

The algorithm written by Ethan and John was already available and run in the standard processing of anti-neutrino events used for this analysis. The code loops over clusters and searches for single strip energy deposits over 20 MeV within a set distance from the vertex, 80 mm both horizontally and vertically. Each strip that meets this goal is counted as a proton candidate.

The physics is to select on the proton's dE/dx curve and associated peak dE/dx.

The peak energy deposit happens at the very end of a proton's travel, where the proton has very little energy left, and is known as the Bragg peak. This is named after William Bragg from his work around 1900 and nowadays is used in proton-based cancer therapies [6]. For the proton this peak occurs right before it stops moving and for the scintillator strips in the the MINERvA detector the peak is found to almost always exceed 20 MeV. The search region is optimized for protons with energy less than or equal to 100 MeV. Such protons typically travel four planes (70 mm) before losing their energy and stopping. Fortunately this is the opposite of what is expected for the majority of neutrons. With the neutron finding algorithm ignoring the vertex region, a proton candidate counted here can not be double counted if ever counting both proton and neutron candidates at the same time. This does not mean there is not a contamination of neutrons being counted as protons because they some times, but not often, interact near the vertex while producing a 20 MeV cluster.

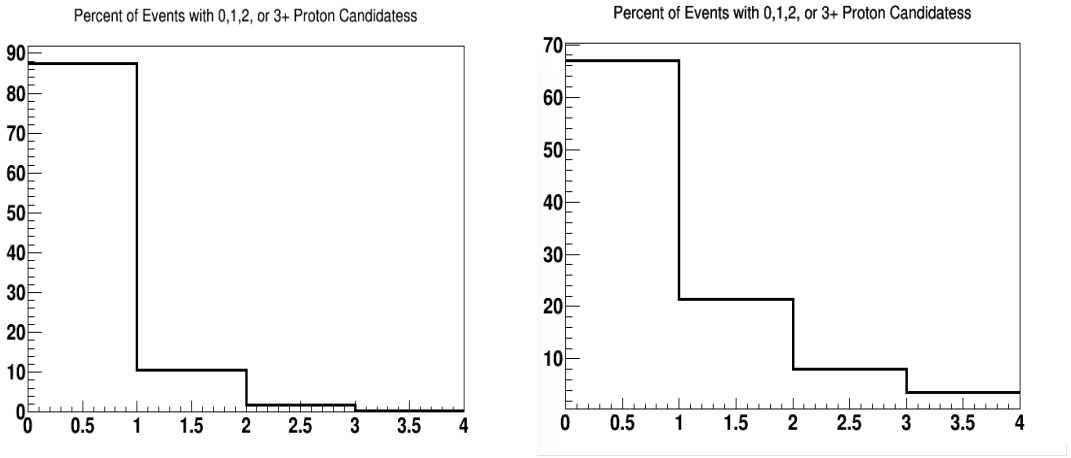

Figure 6.1: Low q3 (left) and mid q3 (right) plots of the percentage of events for simulation with RPA and 2p2h weights being applied that had 1,2,or 3+ proton candidates in them.

Figure 6.1 depicts the percentage of events that had either 1, 2, or 3 plus proton candidates per event. For the low q3 the large majority of the events have no proton candidates available. This makes sense due to the QE interaction type being the predominate of the three interactions for low q3, and only produce a proton if the neutron undergoes a scatter before it exits the nucleus. The predominance of the low q3 QE type interaction sample can be assessed from looking at the reconstructed energy plots shown in Figure 1.2 from Chapter 1.

The mid q3 region is populated more equally than the low q3 by both the QE and delta interactions; nearly a third of the events have protons. The delta interactions described in Chapter 1, which now make up more of the event sample then in the low q3, can have either protons or neutrons in the final state. This is also true of the 2p2h interactions which we see more of in the mid q3.

While protons cause very few, less then 2 percent, of the neutron candidates counted by the neutron counting algorithm, neutrons cause a large portion of the proton candidates in the anti-neutrino sample. Low energy protons are well under-

stood and almost always lose their energy and stop near the vertex before interacting. Due to neutron interactions being governed by a mean free path and an exponential distribution, interactions near the vertex are not rare. Given the anti-neutrino sample has few protons, from Figure 6.1, but many neutrons it is possible the number of neutrons interacting near the vertex is not small in comparison to the number of protons interacting there. This causes the neutron background to the proton candidates to be relatively large.

Candidates in:	Total	QE Region	Dip Region	Delta Region
Low q3	>11.5	22.8	9.5	5.0
Mid q3	> 7.1	23.8	11.3	2.5

Table 6.1: Percentage of proton candidates caused by a neutron in each slice of q3 separated into the total and by hadronic energy region

Figures 6.2 and 6.3 show plots for the number of events with neutron caused proton candidates for the three hadronic energy regions in both the low q3 and mid q3. The plots were made by selecting events that had neutrons plus no other particles and then plotting the number of proton candidates there were per event. Ratios and overall background percentages were then calculated by plotting the total number of proton candidates there were for each of these regions without making any selections on event types. Table 6.1 provides the percentage of the proton candidates in the low q3 and mid q3 that were caused by neutrons broken down into hadronic energy regions. Neutrons caused at least 11.5 percent of the overall proton candidates in the low q3 and 7.1 percent in the mid q3.

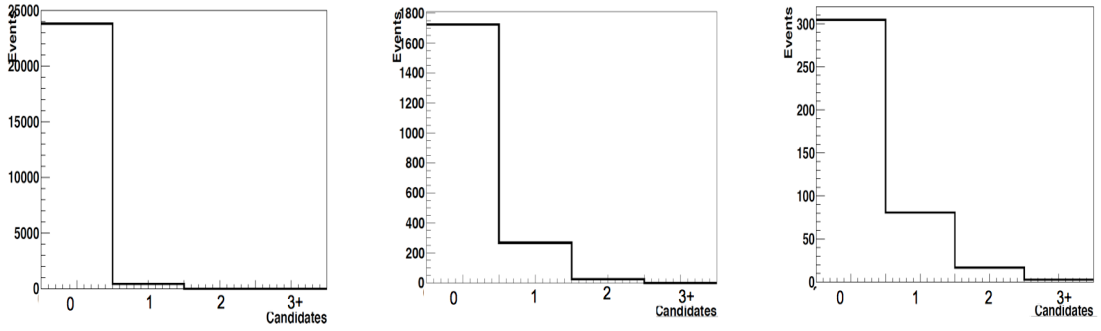

Figure 6.2: Events with a proton candidate, even though there was only one neutron in the simulation. for the QE (left), dip (middle), and delta (right) hadronic energy regions for the low q3.

The neutron background gets smaller as we progress from the QE region, to the dip region, to the delta region, even though there is an increased probability a neutron produces a proton candidate. This is because the probability of finding a proton also increases as we progress through the hadronic energy regions. In the QE region the QE type interaction dominates where only a neutron should be produced with no protons. In the dip and delta regions the 2p2h and delta type interactions

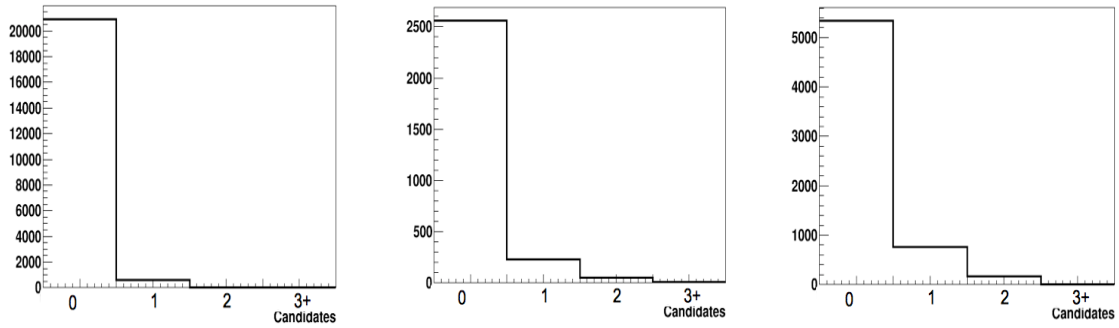

Figure 6.3: Events with neutron caused proton candidates for the QE (left), dip (middle), and delta (right) hadronic energy regions for the mid q3. Events were selected to have neutrons plus nothing else.

are more prominent and both have protons in their final states about half the time. So the neutron-caused proton candidates is large in the QE region, but much smaller fraction in the delta region.

The selected events had only neutrons so this eliminated events that had protons plus neutrons and pions plus neutrons in them that could potentially have had a neutron cause a proton candidate. However, in these events it could not be said what caused the proton candidate so they were ignored. Including these events would only increase the percentage of proton candidates that were caused by neutrons. Thus the 11.5% and 7.1% in Table 6.1 are lower bounds.

The figures seen in the rest of this chapter include fewer plots then those for the neutron counting. The proton tagging algorithm was written without the ability to determine the location of the candidates being counted, because these protons didn't go very far. This made it so that the proton equivalent to the z-distance module plots could not be produced. However the proton candidate plots are available and the hadronic energy plots, while are not different from the those seen in the neutron counting chapter, are put along side them to help with the understanding of each systematic.

6.2 RPA Effect

The addition of the RPA effect has been described as a suppression effect which causes a decrease in the lowest bins in the reconstructed hadronic energy plots repeated below in Figure 6.4. This decrease of events is only applied to the QE interaction rate which rarely produces protons for the anti-neutrino case. So very little change in the proton multiplicity is expected, and a negligible amount is observed.

Plots similar to those seen for the neutron counting in Chapter 4 were also made for counting proton candidates. These plots on the right in Figure 6.4. Here three different simulation models are plotted along with the data in the top portion of the plot with the bottom containing the same models and data but instead their difference from the default simulation which does not contain the RPA nor the 2p2h effects. The ratios are especially helpful to distinguish small differences in the

simulation as each new effect is added in.

The default simulation, which does not contain the RPA nor 2p2h effects, has some differences from the data in proton counting. These disagreements are most noticeable in the ratios of Figure 6.4. As anticipated the number of proton candidates per event does not change significantly with the addition of the RPA weight (shown by the less significant, blue line in the ratio). The small change comes from the decrease in the number of QE events where we already know we should not find a proton which in turn increases the percentage of events where we potentially could. This tweaked simulation with the added RPA weight does not show agreement between data and simulation.

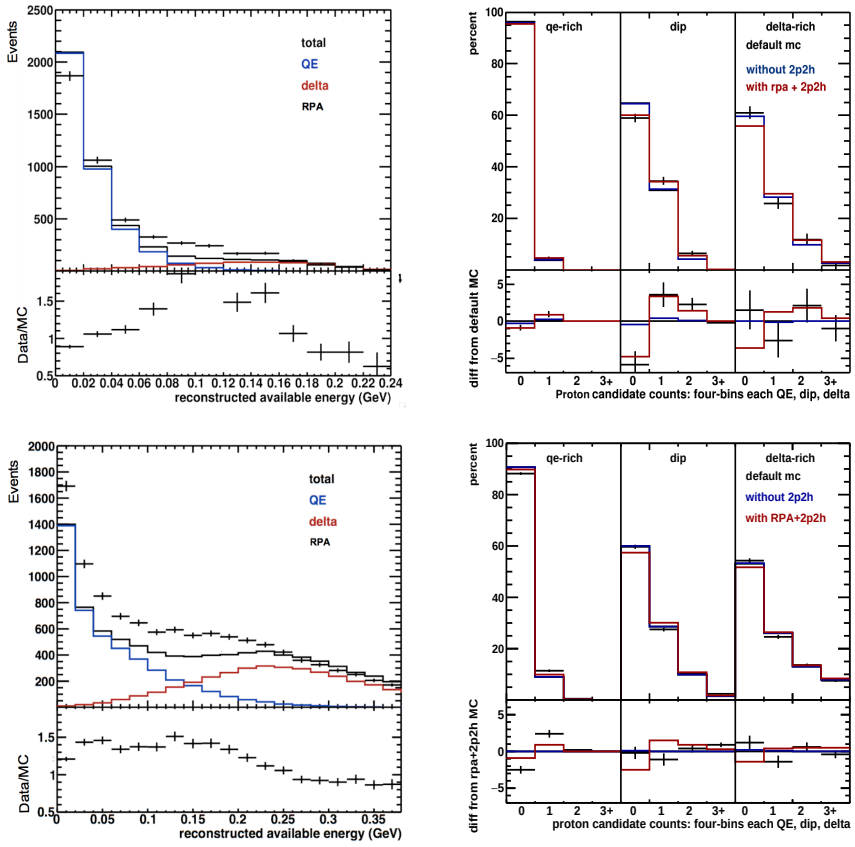

Figure 6.4: Simulation with RPA. Left: Hadronic energy plots for data and simulation where the total (black) is separated into QE (blue), 2p2h(brown), and delta (red). Right: The percentage of events that had 1,2,or 3+ proton candidates in them for each region of hadronic energy: QE, Delta, and Dip. Ratio on bottom taken from the default simulation. Low q3 (top) and Mid q3 (bottom).

6.3 The Valencia 2p2h Model

The addition of 2p2h events causes a significant change in the number of proton candidates found, shown in Figure 6.5. This is because two-particle two-hole events have a proton and a neutron in the final state almost half the time (and two neutrons

slightly more than half). This increase of events with one or more proton describes the data in three of the six regions being considered.

The delta regions and the mid q3 dip region are a different story however. While this increase in total proton candidates per event is beneficial for the low q3, the mid q3 proton counting plots show that the default simulation was already over simulating the number of proton candidates per event in the dip region. This means the addition of events with more protons causes this discrepancy to increase. The same is true for the delta region in both the low q3 and the mid q3 regions.

The magnitudes of these changes are not the same for the low q3 and mid q3. For the low q3 the QE interaction type dominates and the delta interaction contribution is small. This means there are very few protons overall, the few coming from the delta interactions, and so the addition of the 2p2h causes more of a significant change. However, the mid q3 is a mix of the delta and QE interactions thus there are more protons and in turn proton candidates counted before the inclusion of the 2p2h. The addition of the 2p2h here then causes less of a change when compared to the low q3 with the newly added protons.

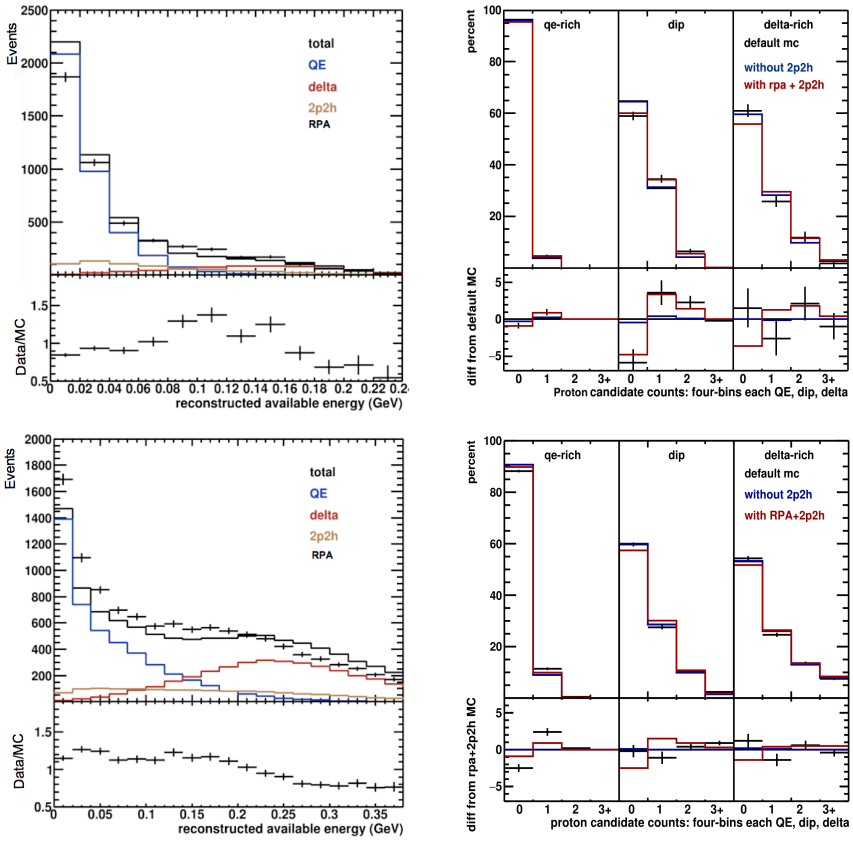

Figure 6.5: Simulation with RPA and 2p2h. Left: Hadronic energy plots for data and simulation where the total (black) is separated into QE (blue), 2p2h(brown), and delta (red). Right: The percentage of events that had 1,2,or 3+ proton candidates in them for each region of hadronic energy: QE, Delta, and Dip. Ratio on bottom taken from the default simulation. Low q3 (top) and Mid q3 (bottom).

6.4 Birks Systematic

Birks' Law is a formula that quantifies the light yield per path length as a function of the energy loss per path length for a particle traversing a scintillator. This parameter is one of the largest contributors of uncertainty within proton counting as shown by the neutrino MEC analysis [7]. This was seen not only in the neutrino case but in the anti-neutrino simulation as well.

The MINERvA test beam used stopping protons to measure the Birks' suppression for the scintillator in the MINERvA detector. The uncertainties on that measurement are exactly what apply to this situation. The MC energy deposits are recomputed using the one-sigma varied Birks' parameter, and the 20 MeV threshold proton counts are redone. More Birks' suppression will decrease the proton count total and the decrease in the Birks' suppression should cause an increase in the proton counts total. The Birks' suppression increase plots are shown on the left of Figure 6.6 while the Birks' suppression decrease plots are shown on the right of Figure 6.6.

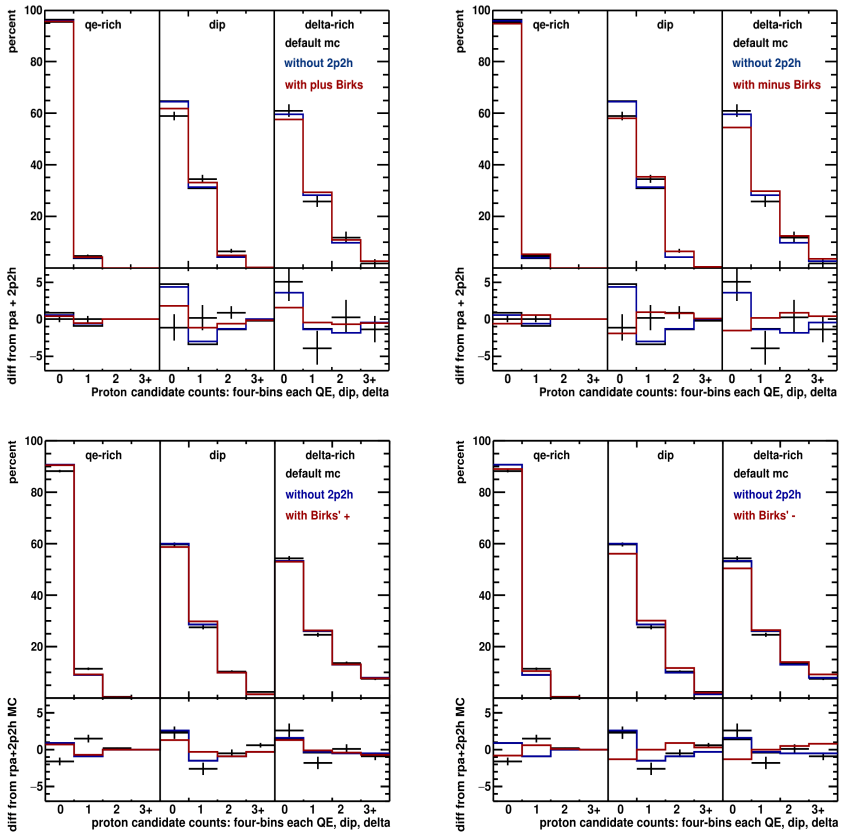

Figure 6.6: Simulation includes RPA, 2p2h,and increase (left) or decrease (right) in Birks Suppression the percentage of events that had 1,2,or 3+ proton candidates in them for each region of hadronic energy: QE, Delta, and Dip. Ratio on bottom taken from the simulation with RPA and 2p2h. Low q3 (top) and mid q3 (bottom).

The decrease of the Birk's suppression gives an overall increase in the total

number of proton candidates per event in all three regions of hadronic energy for both the low q3 and the mid q3 (red line on the right plots). The increase in each region is slightly different from one another mainly due to the number of protons typically found in each of them. The QE hadronic energy regions, for both the low q3 and the mid q3 tend to have the smallest increase of candidates per event. This makes sense due to this region having mainly QE type interactions, where protons shouldn't appear, and a small number of 2p2h and delta type interactions. The magnitude of the change is slightly greater in the mid q3 QE hadronic energy region due to there being a higher contamination from the other interaction types allowing for more proton candidates to be counted. The dip and delta regions within the low q3 or mid q3 are similar in magnitude and between three-momentum transfer slices as well.

Increasing the Birks' suppression has exactly the opposite effect, and has a magnitude almost as large as turning off the 2p2h component. The effect is larger for the dip and delta regions in both the low q3 and the mid q3 then it was for the QE hadronic energy region. There are now many more events with clusters that just make the 20 MeV cut then there were before allowing for there to be more protons per event.

6.5 Hadronic Energy Scale Uncertainty

The MINERvA test beam data constrain the uncertainty on the proton and pion response to 4%. Because neutron response was not constrained by the test beam, a larger uncertainty of 10% is considered. Adjusting the hadronic energy to compensate for this uncertainty for the events being studied in this analysis should have an impact on the number of proton candidate there are per event just as it did for the number of neutron candidates because once again events will be migrated from one slice of momentum-transfer to another and from one region of hadronic energy to another.

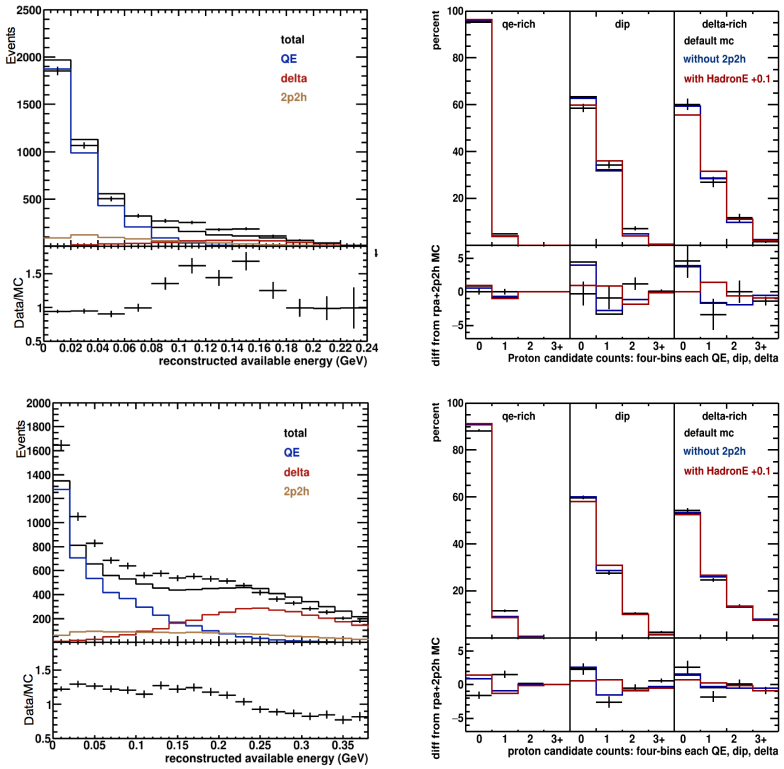

Figure 6.7: Simulation with RPA, 2p2h, and Hadron Energy 10 percent increase. Left: Hadronic energy plots for data and simulation where the total (black) is separated into QE (blue), 2p2h(brown), and delta (red). Right: The percentage of events that had 1,2,or 3+ proton candidates in them for each region of hadronic energy: QE, Delta, and Dip. Ratio on bottom taken from the simulation with RPA and 2p2h. Low q3 (top) and Mid q3 (bottom).

Figure 6.7 shows the proton candidate counting plots and the associated hadronic energy for the increase in the hadronic energy by ten percent. This increase should, like in the neutron case, migrate events with protons to higher slices of q3 and from the QE hadronic energy region into the dip region and from the dip region into the delta region. The events without protons in them should have smaller amounts of hadronic energy so the increase of four percent should not be enough to move these events from the regions they already live in. This means events with protons are leaving and those without, in the zero candidate bin, are staying. The movement from low q3 into mid q3 and then from mid q3 to higher q3 are not symmetric and neither is the movements between hadronic energy regions. Thus the percentage of events without proton candidates increase and those with decrease which brings simulation away from agreement with the data in most instances.

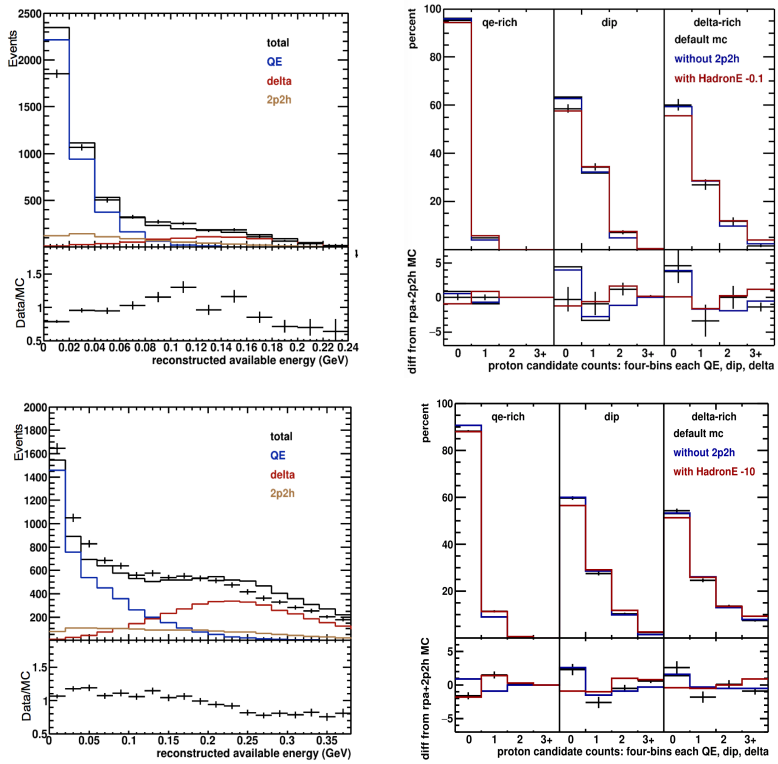

Figure 6.8: Simulation with RPA, 2p2h, and Hadron Energy 10 percent decrease. Left: Hadronic energy plots for data and simulation where the total (black) is separated into QE (blue), 2p2h(brown), and delta (red). Right: The percentage of events that had 1,2,or 3+ proton candidates in them for each region of hadronic energy: QE, Delta, and Dip. Ratio on bottom taken from the simulation with RPA and 2p2h. Low q3 (top) and Mid q3 (bottom).

Decreasing the hadronic energy by ten percent has an equal and opposite effect on the proton counting plots then the increase of ten percent. Figure 6.8 shows the proton candidate plots along with the hadronic energy plots for this systematic. Here we see a decrease in the percentage of events without protons and an increase in the percentage of events with protons. This is once again due to the events with protons moving, now in the opposite direction towards lower q3 and hadronic energy regions, and those without protons staying nearly in place. This systematic brings the simulation in the mid q3 QE region closer to agreement with the data. In the QE-rich region, the magnitude is similar in size to adding 2p2h events. In these distributions, lowering the hadronic energy scale seemed to describe the data, opposite of the neutron case. This contradiction indicates other effects are playing a significant role.

The increase in hadronic energy by ten percent gives an effect similar and just over double in magnitude to the proton candidate count in comparison to the increase by four percent. Once again the change of four percent has been omitted.

6.6 Increase Pion Absorption for Delta Production

The systematic implements an increase, about double, in the number of events with the product pion of a delta interaction being absorbed by the nucleus with a second nucleon kicked out instead of a single nucleon and a pion. This increase is coupled by a decrease, nearly 20 percent, in events where the product pion of the delta interaction does make it out of the nucleus such that the overall delta interaction rate does not change. This increase in events with two nucleons in the final state should appear very similar to the addition of the the 2p2h events (and in these plots will have the opposite effect as removing 2p2h events), due to these events having two nucleons in the final state, and thus any changes in proton counting should be visible near the dip and delta defined hadronic energy regions. The proton counting plots and associated hadronic energy plots for this systematic can be seen in Figure 6.9.

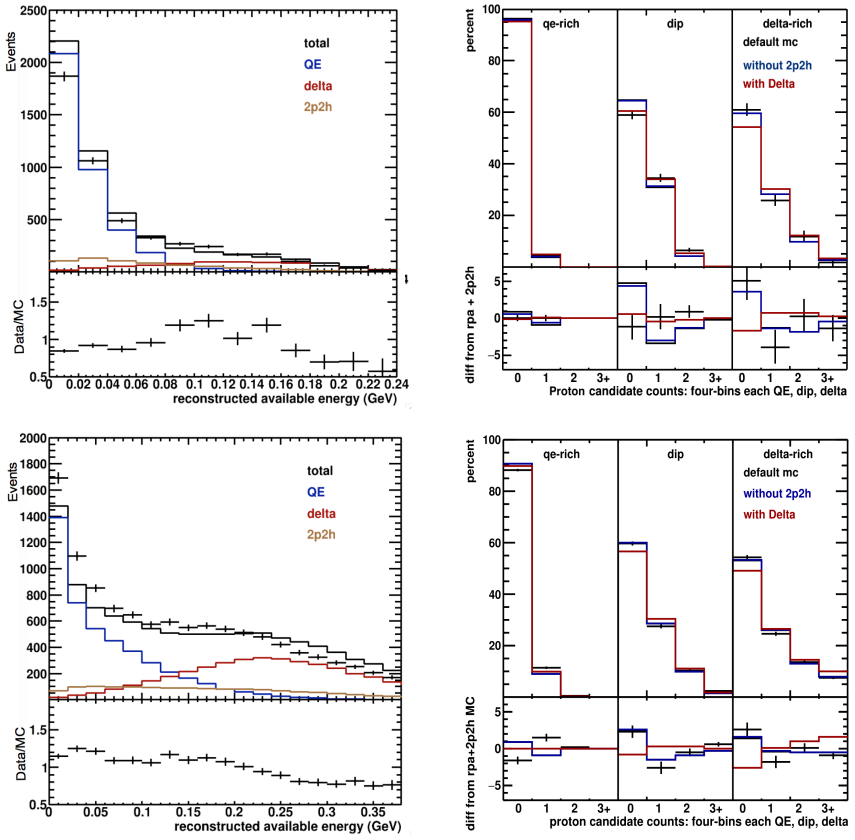

Figure 6.9: Simulation with RPA, 2p2h, and Delta. Left: Hadronic energy plots for data and simulation where the total (black) is separated into QE (blue), 2p2h(brown), and delta (red). Right: The percentage of events that had 1,2,or 3+ proton candidates in them for each region of hadronic energy: QE, Delta, and Dip. Ratio on bottom taken from the simulation with RPA and 2p2h. Low q3 (top) and Mid q3 (bottom).

The increase of events with two nucleons, some being protons and other neutrons, but a same number of total events causes an increase in numbers of protons per event. The QE-rich hadronic energy regions for both the low q3 and the mid q3 have very few delta events within for the default simulation so they remain untouched.

The low q3 shows an increase of proton candidates per event in the delta region and decrease of proton candidates per event for the dip region. The speculation here is that the added delta interactions that appear in the delta hadronic energy region are those with protons in the final state because they are typically more visible and more energetic then neutrons. The interactions that appear in the dip region are those with neutrons in the final state which are less visible. The mid q3 proton counting shows an increase of proton candidates in both the dip and delta hadronic energy regions. This could be due to the neutron count for the entire mid q3 appearing very similar to the dip and delta regions because the mid q3 is a mix of the QE and delta type interactions.

6.7 Increase the QE Axial Mass Parameter

The QE weight applied here causes an increase in the number of QE type interactions as a function of the three-momentum transfer. The higher the three-momentum transfer the more the QE interaction type is increased meaning the mid q3 in affected more then the low q3. This can be seen when looking at the hadronic energy plots in Figure 6.10 and comparing to the previous hadronic energy plots without QE in the last section in Figure 6.5. This total increase of QE events is somewhat proportional throughout the hadronic energy plots for either the low q3 or mid q3, meaning proportional percentages of events are added to each bin. The effects of this addition on counted proton candidates can be seen in Figure 6.10.

The increase of the QE interaction type has little effect on the proton counts for the low q3 and mid q3. The low q3 has been described as being dominated by the QE interaction type and having few proton candidates to be counted so the increase of these events does not change the proton multiplicity. The mid q3 proton counting plot shows a bigger, but still small increase in the number of events without a proton candidate in it as the QE systematic is applied. This is due to the mid q3 being a mix of QE, dip, and delta events instead of almost all QE.

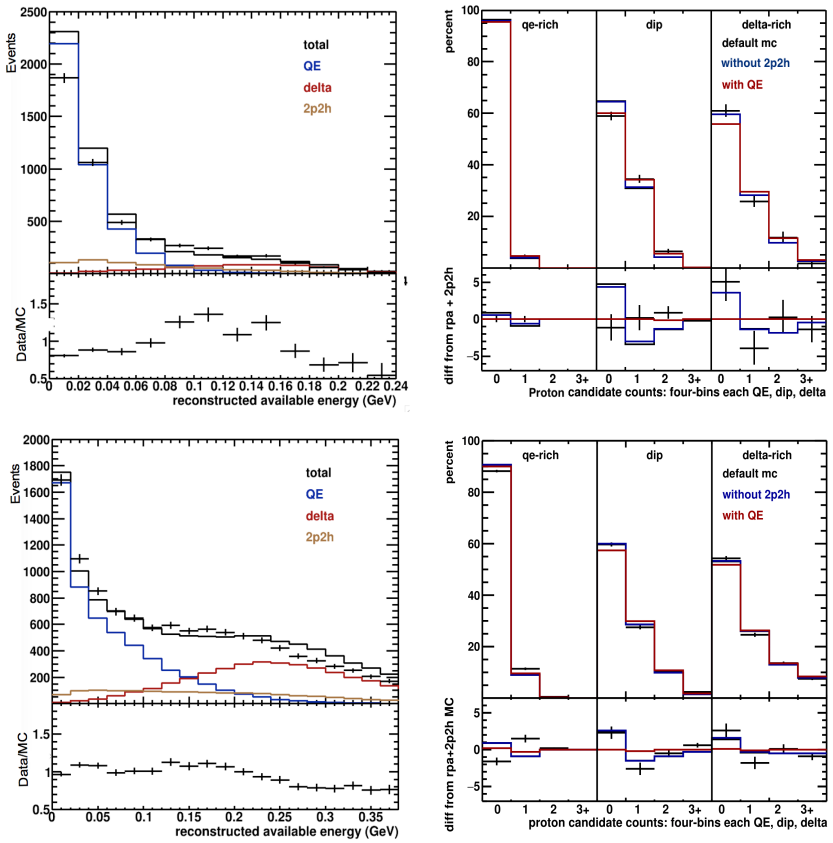

Figure 6.10: Simulation with RPA, 2p2h, and QE. Left: Hadronic energy plots for data and simulation where the total (black) is separated into QE (blue), 2p2h(brown), and delta (red). Right: The percentage of events that had 1,2,or 3+ proton candidates in them for each region of hadronic energy: QE, Delta, and Dip. Ratio on bottom taken from the simulation with RPA and 2p2h. Low q3 (top) and Mid q3 (bottom).

6.8 Flux Systematic

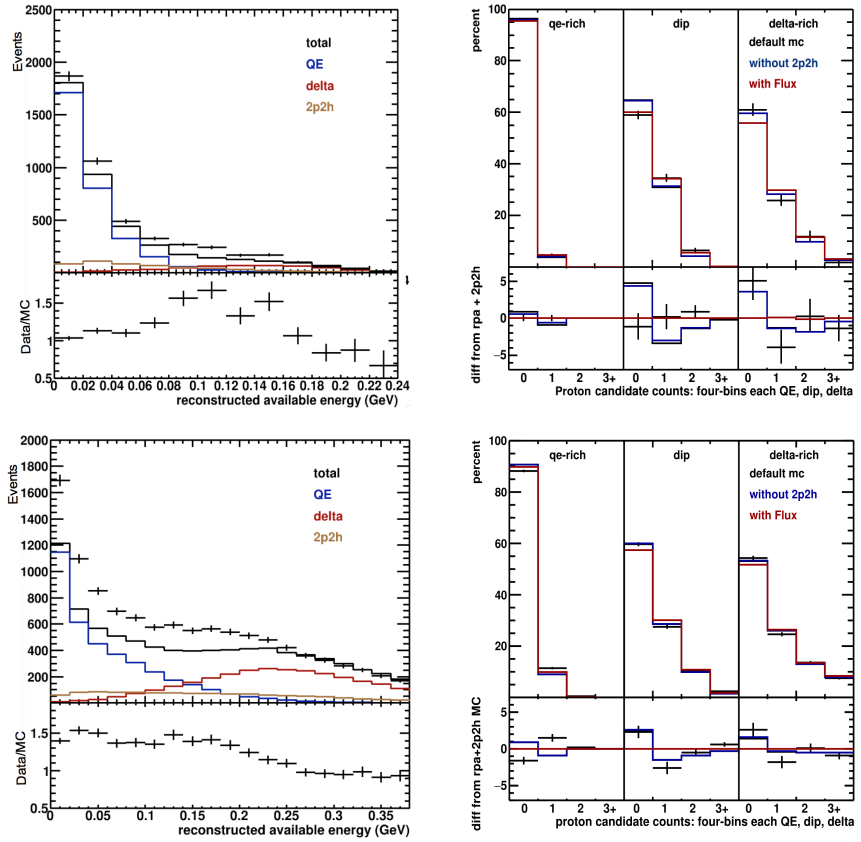

Figure 6.11: Simulation with RPA, 2p2h, and Flux. Left: Hadronic energy plots for data and simulation where the total (black) is separated into QE (blue), 2p2h(brown), and delta (red). Right: The percentage of events that had 1,2,or 3+ proton candidates in them for each region of hadronic energy: QE, Delta, and Dip. Ratio on bottom taken from the simulation with RPA and 2p2h. Low q3 (top) and Mid q3 (bottom).

The flux weight should have a similar effect on the proton counting as it did on the neutron counting in that it shouldn't cause a change. The flux should only change the overall event rate and should do so proportionally for each interaction type so that the overall percentages of each do not change. This lack in change of percentages between interaction type should then cause the proton counting to remain untouched and as seen in Figure 6.11.

6.9 Intra-nuclear Re-scattering Mean Free Path

This weight changes the mean free path for nucleons and pions that re-scatter before leaving the nucleus causing the final states we see in the detector to change. Increasing the nucleon MFP decreases the likelihood nucleons interacting inside of

the nucleus. (Changing the nucleon mean free path is separate from changing the pion mean free path.) The same logic can be used for the other three variations of this systematic. The increase of the nucleon mean free path is shown in Figure 6.12.

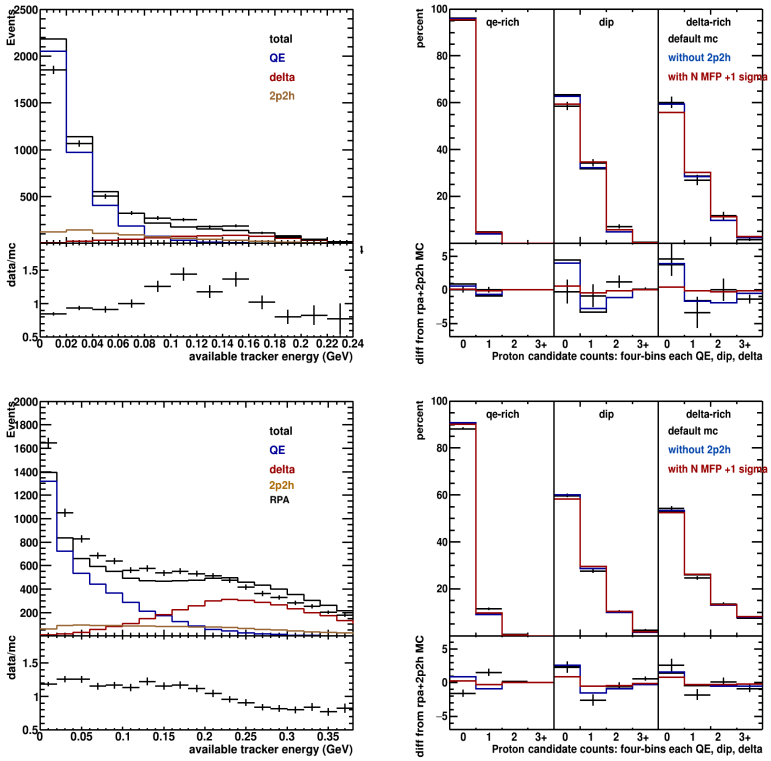

Figure 6.12: Simulation with RPA, 2p2h, and Nucleon Mean Free Path increase. Left: Hadronic energy plots for data and simulation where the total (black) is separated into QE (blue), 2p2h(brown), and delta (red). Right: The percentage of events that had 1,2,or 3+ proton candidates in them for each region of hadronic energy: QE, Delta, and Dip. Ratio on bottom taken from the simulation with RPA and 2p2h. Low q3 (top) and Mid q3 (bottom).

Increasing the nucleon mean free path causes some change in the proton counting plots. There is an increase in the events without proton candidates in them because fewer neutrons re-scattered or kicked out protons. The effect on this distribution is more significant than it is for the neutron multiplicity distribution.

Decreasing the nucleon mean free path gives the opposite but a larger effect. The increase of the mean free path causes the intra-nuclear interactions to increase. If neutrons are interacting within the nucleus it is now possible for a proton to exit the nucleus instead. Because there are so many neutrons in the beginning and the likelihood of a neutron or a proton leaving after every neutron's intra-nuclear reaction are equal then the proton count must increase.

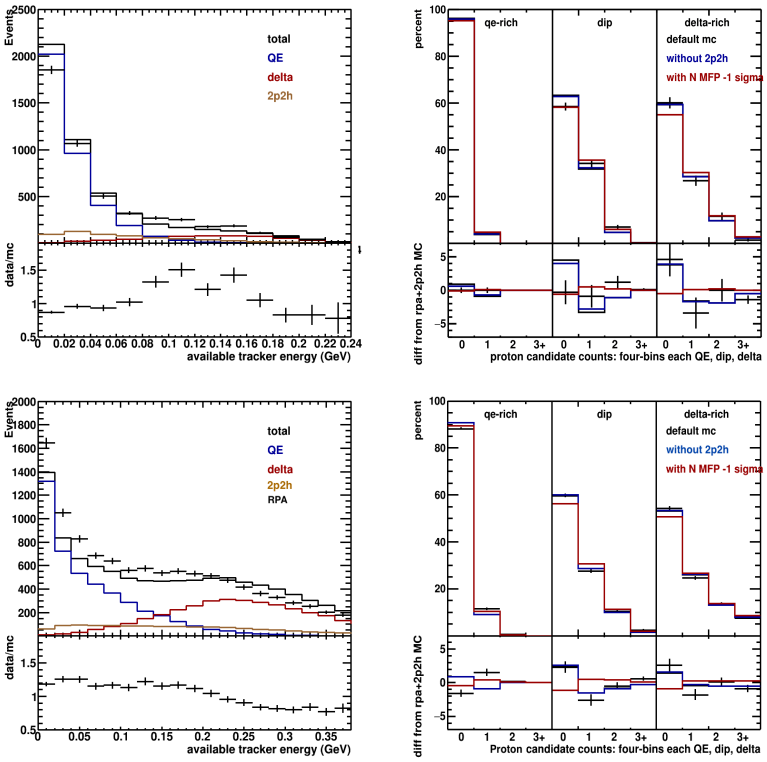

Figure 6.13: Simulation with RPA, 2p2h, and Nucleon Mean Free Path decrease. Left: Hadronic energy plots for data and simulation where the total (black) is separated into QE (blue), 2p2h(brown), and delta (red). Right: The percentage of events that had 1,2,or 3+ proton candidates in them for each region of hadronic energy: QE, Delta, and Dip. Ratio on bottom taken from the simulation with RPA and 2p2h. Low q3 (top) and Mid q3 (bottom).

The change of the pion's intra-nuclear mean free path has no noticeable affect on the proton counting plots. The increase or decrease of this mean free path is only by one sigma and if the number of events with pions re-scattering in the nucleus is also small, with there being a chance for either a proton or neutron to leave the nucleus instead of or with the pion, then the increase or decrease of protons alone should also be very small.

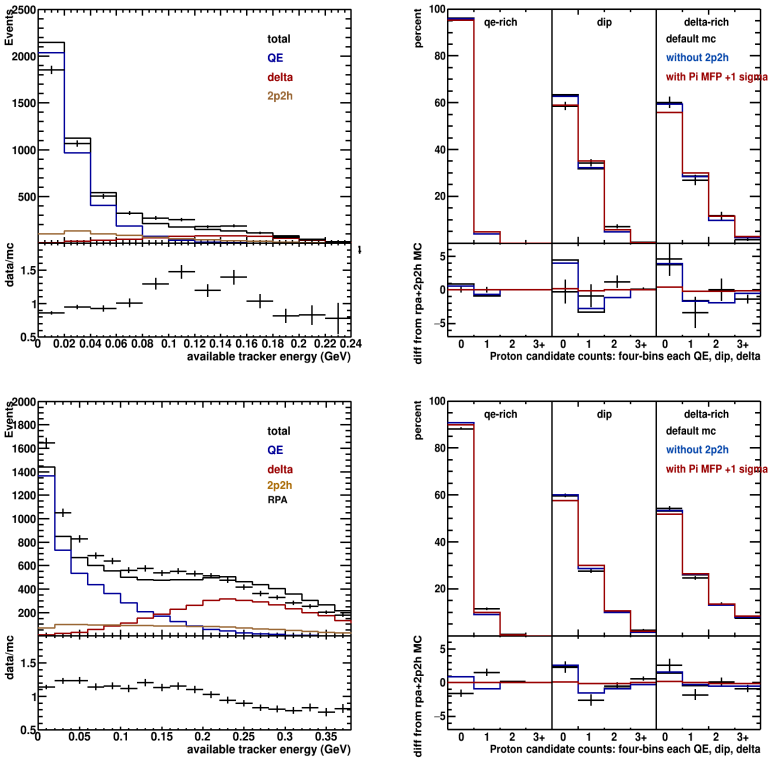

Figure 6.14: Simulation with RPA, 2p2h, and Pion Mean Free Path increase. Left: Hadronic energy plots for data and simulation where the total (black) is separated into QE (blue), 2p2h(brown), and delta (red). Right: The percentage of events that had 1,2,or 3+ proton candidates in them for each region of hadronic energy: QE, Delta, and Dip. Ratio on bottom taken from the simulation with RPA and 2p2h. Low q3 (top) and Mid q3 (bottom).

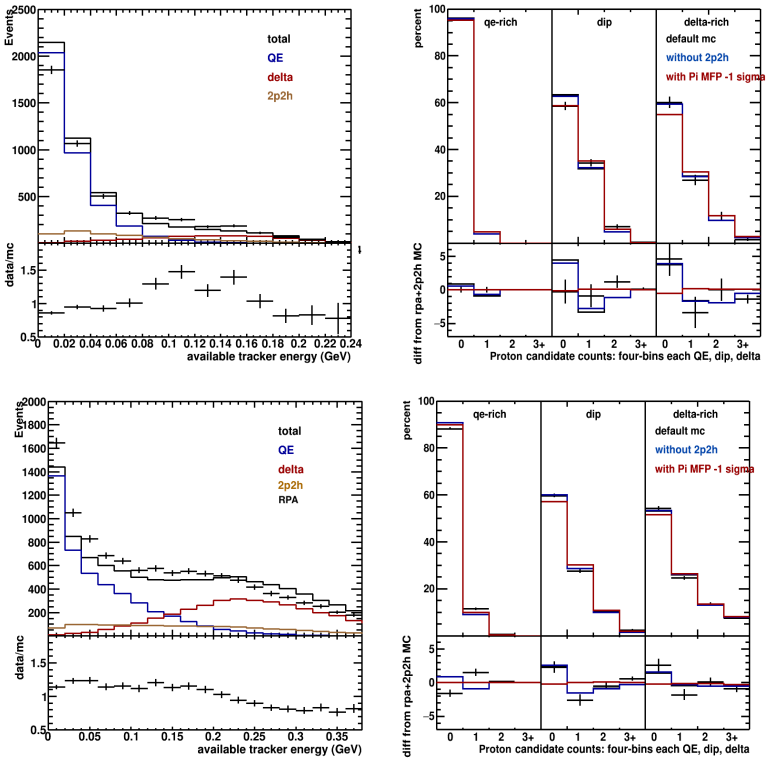

Figure 6.15: Simulation with RPA, 2p2h, and Pion Mean Free Path decrease. Left: Hadronic energy plots for data and simulation where the total (black) is separated into QE (blue), 2p2h(brown), and delta (red). Right: The percentage of events that had 1,2,or 3+ proton candidates in them for each region of hadronic energy: QE, Delta, and Dip. Ratio on bottom taken from the simulation with RPA and 2p2h. Low q3 (top) and Mid q3 (bottom).

Chapter 7

Additional Studies and Uncertainties

7.1 Tuned 2p2h and Uncertainties

During the end of this analysis a new version of the 2p2h model being used was released. The new version is referred to as the Phil Tuned 2p2h. This version of the 2p2h uses a a 2D Gaussian fit in true q0 and q3 to give events new weights which increases events in the dip region of the hadronic available energy plots. Using this new 2p2h, neutron counting was completed for the base set seen in Figure 7.1. This base set includes the default data, turning on the RPA effect, and then turning on the new tuned 2p2h. These new neutron counting plots and hadronic energy plots also display the error bars which include the errors taken in quadrature for the complete set of MINERvA systematics, not just the important ones studied in the previous chapters. These new error bars help determine whether or not the simulation is within agreement of the data given the uncertainties of each of the systematics, some of which were not studied in this analysis because they were expected to have small or negligible uncertainties.

The full set of systematics can be ranked by importance. The largest and average difference between the data and the tweaked simulation for each systematic are determined using the neutron and proton counting in each of the 12 bins. For neutron counting, the top three systematics that have both the largest difference and largest average difference in order from largest to smallest are the addition the ten percent hadronic energy change, the tuned 2p2h, and the weighting down of the neutron cross section using the hadron re-weight tool. For the proton counting case the top three are the addition of the tuned 2p2h, the ten percent hadronic energy change, and the Birks' suppression change. These largest uncertainties were all studied for this analysis.

The tuned 2p2h is done by fitting the neutrino data set, where it fills in the dip region and gives good agreement there. The resulting parameters become a prediction for this anti-neutrino analysis, and also gives better agreement for both slices of q3. The addition of this 2p2h does not cause a significant change in the neutron counting plots. The over simulation of events with candidates in the default

simulation still exists so the addition of the RPA and the now tuned 2p2h pushes this agreement further apart. The z-distribution plots also have the over simulation of candidates in the forward direction as before.

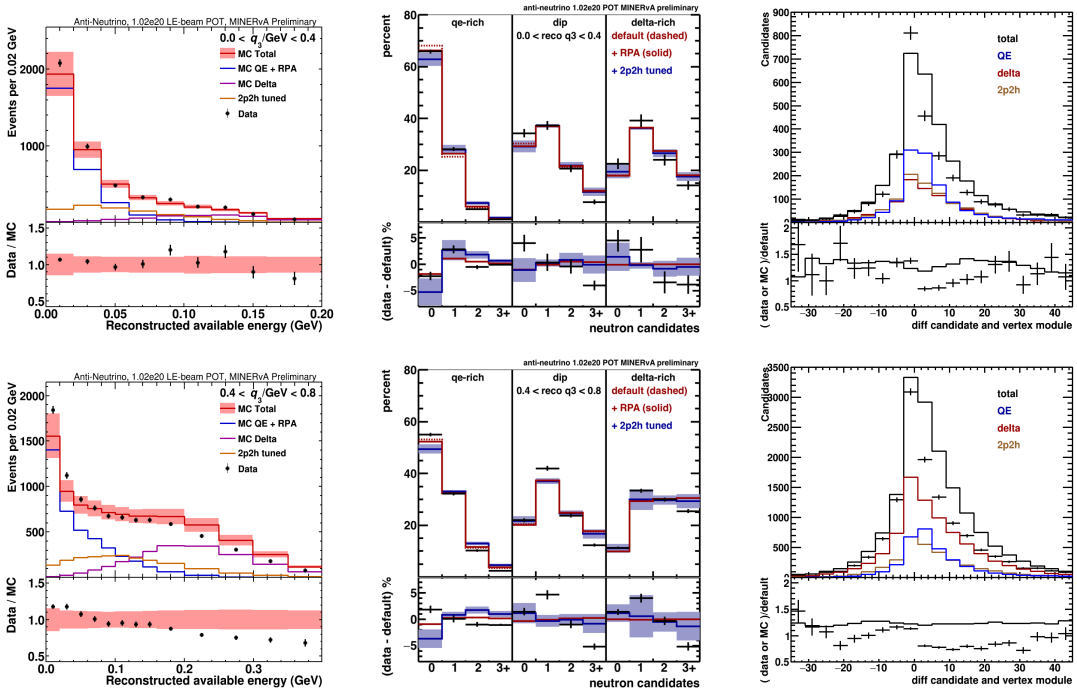

Figure 7.1: Simulation with RPA and tuned 2p2h. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and RPA plus tuned 2p2h. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), tuned 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the default simulation while the ratio for the left plots is to the current simulation. The error bars seen are for the errors taken in quadrature for each of the systematics imposed on top of the RPA+2p2h model tweaks.

The X^2 calculations for the neutron counting have also been completed. These X^2 values are found for the addition of each systematic or model tweak to determine whether or not these changes are an improvement to the model or not. The X^2 is calculated by summing the ratio between square of the difference of the data and simulation values and their uncertainties in each bin for neutron counting. This is seen in Equation 7.1 below.

$$X^2 = \Sigma \frac{(Data - Sim)^2}{\sigma^2} \tag{7.1}$$

The table of numbers shown in Table 7.1 displays the X^2 values for both the proton and neutron counting in each slice of q3. These X^2 values actually include the full systematic co-variance, reusing code from the 2016 proton counting result. The values for the total, the three hadronic energy regions, and the combination of

the QE-rich and dip regions are given. The proton counting values are calculated using only ten bins instead of twelve, combining four from each of the three hadronic energy regions, because the 2 and 3+ candidate bins in the qe-rich region for the proton counting have few to no events falling within. It is obvious from these chi squares that there is no distinct pattern and the addition of these model tweaks seems to given an improvement in some regions but not in others.

			Components			
		Total X2	qe-rich	dip	qe+dip	delta-rich
Neutron	Pion	24.0016	11.898	7.908	19.807	4.195
Low q3	RPA	18.7624	7.378	8.975	16.352	2.410
12 bins	2p2h	17.442	9.305	8.311	17.615	-0.173
	Phil	18.8275	16.639	3.197	19.835	-1.008
Neutron	Pion	32.4329	-5.041	33.974	28.933	3.500
Mid q3	RPA	29.3127	-4.428	32.562	28.134	1.179
12 bins	2p2h	33.5198	4.071	32.841	36.912	-3.393
	Phil	46.9156	31.658	14.980	46.638	0.277
Proton	Pion	9.73987	0.149	4.813	4.962	4.778
Low q3	RPA	10.0168	-0.584	5.933	5.349	4.668
10 bins	2p2h	10.7576	0.054	1.655	1.708	9.049
	Phil	11.0882	0.528	0.613	1.140	9.948
Proton	Pion	11.3038	4.312	-1.088	3.224	8.080
Mid q3	RPA	11.4164	4.131	-1.040	3.090	8.326
10 bins	2p2h	9.37396	2.289	0.147	2.435	6.939
	Phil	14.9485	0.615	14.485	15.100	-0.151
			(2 bins proton)			

Table 7.1: X^2 calculations for both neutron and proton counting. The values for each slice of momentum-transfer are given and broken down into the hadronic energy regions defined.

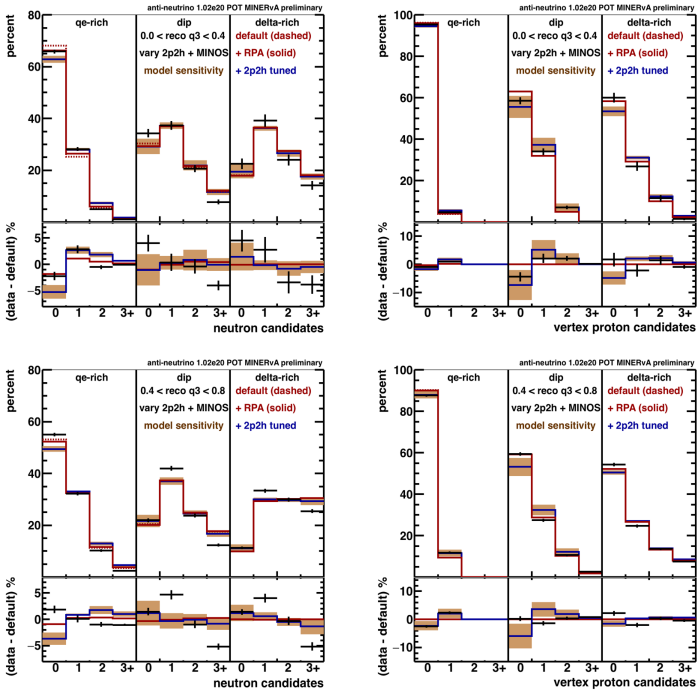

Figure 7.2: Simulation with RPA and tuned 2p2h. Plots for the low q3 (top) and mid q3 (bottom). Neutron (left) and proton (right) counting plots include data, the default simulation, default plus RPA, and RPA plus tuned 2p2h. Ratio is to the default simulation. The error bars seen are the errors taken in quadrature for the 2p2h variations and a MINOS resonance suppression model.

There are three variations to the new tuned 2p2h which are not discussed in this thesis but highlight the uncertainties in the model. These variations include choosing to favor a specific initial state for the 2p2h model which in turn favors either the n-n or p-n final states. This causes a large uncertainty in the number of neutron and proton candidates. The final variation causes an increase of events that appear in the QE-rich hadronic energy region. The plots in Figure 7.2 are both the base neutron and proton counting plots but now show the systematic error bars associated with these 2p2h variations and also include an error from a MINOS "RPA-like" resonance suppression variation in relation to the default tune. These error bars help identify where the simulation may be sensitive to the addition of the 2p2h model.

7.2 Neutron HP - Neutron Counting

A Geant4 variation nicknamed NeutronHP has a more precise modeling of how lower energy neutrons propagate and interaction within the MINERvA detector, which includes a majority of the neutrons found in this analysis. Completing neutron counting using this Geant4 variation will help identify systematic uncertainties with the default Geant4.

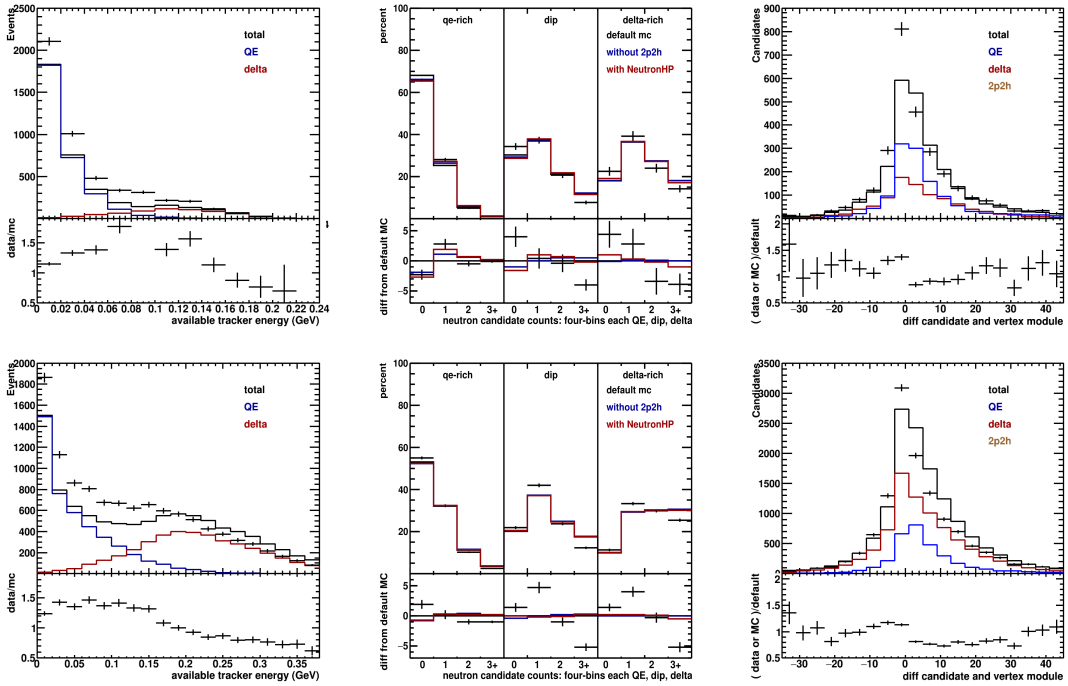

Figure 7.3: Simulation with default Neutron HP, Neutron HP plus RPA, and default GENIE. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the simulation with RPA and 2p2h while the ratio for the left plots is to the default Neutron HP.

The distributions displayed in 7.3 are similar to the previous sets found in this chapter. The z-distribution and hadronic energy plots are still plotted in the same way with the current simulation information being plotted along with data. The neutron counting plot now includes the default GENIE and Geant4, default GENIE and Geant4 with RPA turned on, and then the default GENIE with Geant4 NeutronHP with RPA turned on. The NeutronHP variation was generated without 2p2h events, so the RPA only comparison is the best, consistent one we can make.

The hadronic energy plots are somewhat different then what is seen in previous sections. The x-axis has now been scaled by 17% for aesthetic purposes. There are more discrepancies in the dip and delta regions using the Geant4 NeutronHP variation than was seen in the default Geant4. The 2p2h is not modeled so the addition of this would help the discrepancy of the under simulation of events in between the QE and delta interactions seen in the hadronic energy plots. The addition of the Geant4 NeutronHP to the simulation causes small changes, not as large as the biggest systematics like the addition of the 2p2h or the 10 percent hadronic energy change, for the low q3 and almost no change for the mid q3. The NeutronHP pulls simulation in the direction of the data for the low q3 qe-rich and delta-rich regions but in the opposite direction in the dip region.

The z-distribution plots for the Neutron HP variation with RPA have more candidates in them than found for the default GENIE and Geant4 with RPA. The low q3 describes data well in bins not near the vertex (in the middle of the histogram), where there is an under simulation of candidates. The mid q3 has an over simulation of candidates in the forward direction similarly to what the default GENIE and Geant4 with RPA simulation had. The backward candidates for the mid q3 describe data well.

The addition of a 2p2h process would increase candidates for both slices of q3. The increase of candidates in the mid q3 would increase the over simulation that is already present. For the low q3 a majority of these candidates would appear in the middle few bins causing better agreement between simulation and data.

7.3 NeutronHP - Proton Counting

The NeutronHP variation uses better predictions of how low energy neutrons interact in comparison to the default Geant4 and is useful for proton counting because there is a neutron background to the proton counts in this sample. The proton counting plots are different from the previous set with default GENIE and Geant4, default GENIE and Geant4 plus RPA, and default GENIE and Geant4 with RPA and 2p2h in that the final set is now default GENIE plus NeutronHP Geant4 with RPA.

Turning on the NeutronHP Geant4 causes tiny changes in comparison to the default Geant4. The low q3 qe-rich and dip regions both have a small increase in the percentage of events with candidates in them, pull the qe-rich towards data and the dip region away from data. The delta-rich region sees a significant decrease in the percentage of events with candidates pulling the simulation towards the data here. The mid q3 sees almost no change between the two Geant4 variations except in the qe-rich region where there is now slightly more events with candidates. The effect of the NeutronHP addition is smaller then that of adding in a 2p2h process.

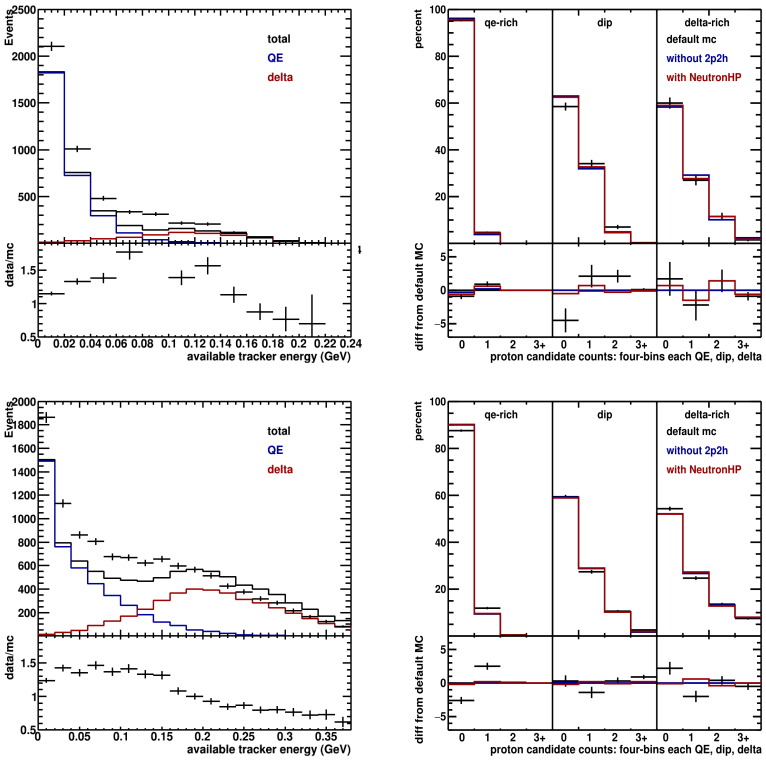

Figure 7.4: Simulation with default GENIE and Geant4, default GENIE and Geant4 with RPA, and default GENIE and Geant4 NeutronHP with RPA. Left: Hadronic energy plots for data and simulation where the total (black) is separated into QE (blue), 2p2h(brown), and delta (red). Right: The percentage of events that had 1,2,or 3+ proton candidates in them for each region of hadronic energy: QE, Delta, and Dip. Ratio on bottom taken from the default GENIE and Geant4 simulation. Low q3 (top) and Mid q3 (bottom).

The addition of a 2p2h process would increase the percentage of events with candidates for all three hadronic energy regions in both slices of q3. This would be beneficial for both the qe-rich and dip regions in the low q3 and the qe-rich of the mid q3. The dip and delta-rich regions in the mid q3 already have a larger percentage of events with candidates then the data so the 2p2h would increase these discrepancies. It is hard to say whether or not this would be helpful for the delta-rich region in the low q3. This addition would likely pull the simulation from agreement in the first two bins of the low q3 delta-rich region.

7.4 Candidate Threshold Changes

It has been demonstrated there is an over simulation of neutron candidates and percentage of events with neutron candidates in them throughout this thesis. We can determine roughly how many candidates need to be removed from the simulation in order to better describe data. There are two ways. One is to force a higher threshold for the MC than the data (a 2 MeV candidate threshold, on top of the

cluster threshold of 1.5 MeV). A second is to eliminate a fraction, here either 12% or 24% of candidates under 10 MeV are eliminated. A third is to eliminate a fraction, but only when the MC says a neutron actually caused the candidate. The following plots shown in Figure 7.5 are the base set and have been repeated for easier comparisons.

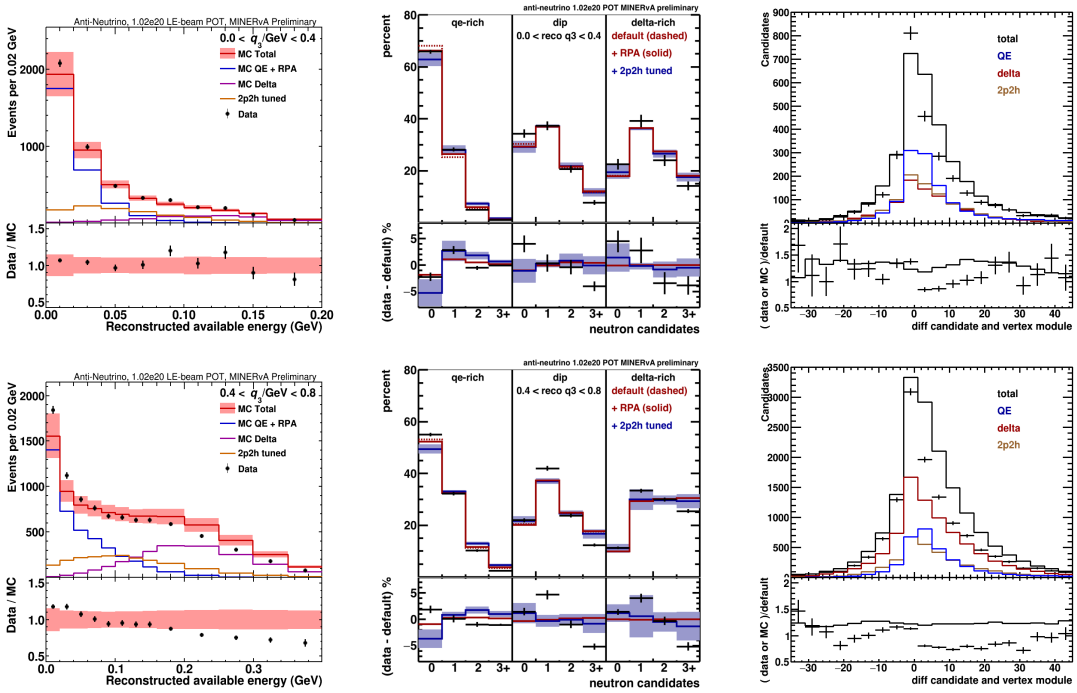

Figure 7.5: Simulation with RPA and tuned 2p2h. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and RPA plus tuned 2p2h. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), tuned 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the default simulation while the ratio for the left plots is to the current simulation. The error bars seen are for the errors taken in quadrature for each of the systematics imposed on top of the RPA+2p2h model tweaks.

The first set of plots in Figure 7.6 show the changes to the neutron counting and z-distribution plots when the 2 MeV candidate threshold is included. The neutron counting plot shows the 2 MeV cut included for each variation of the simulation but not for the data, where candidates are made with the default algorithm. The z-distribution plot is the same as before except now the ratio to the default means the ratio to the default with the 2 MeV candidate energy threshold cut being applied. The hadronic energy plots will not change because these do not depend on the neutron candidates being counted.

The addition of the 2 MeV candidate cut decreases the number of candidates there are overall and helps the simulation describe the data in neutron counting. Both the low q3 and the mid q3 describe data well when the RPA and 2p2h is

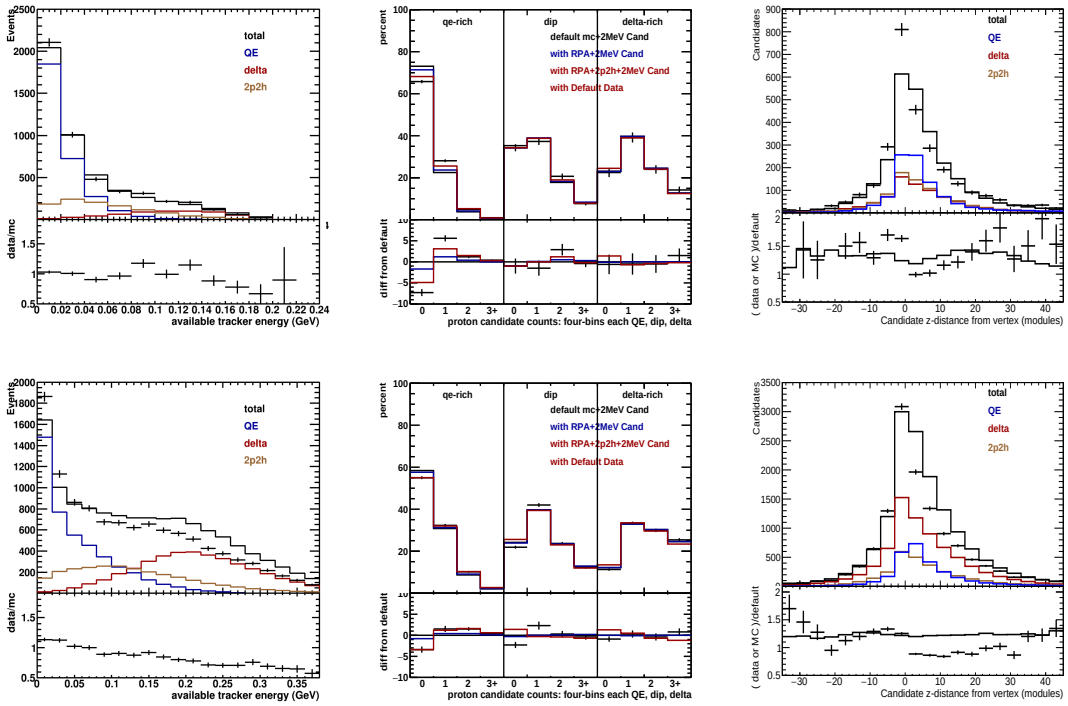

Figure 7.6: Simulation with RPA, 2p2h, and 2 MeV candidate cut and data with default algorithm. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the default simulation with a 2 MeV candidate energy threshold while the ratio for the left plots is to the current simulation.

turned on with the exception of a few bins considering statistical error bars only. With systematic error bars included all bins would describe data. These error bars can be seen in Figure 7.1 in the final section of Chapter 5.

The z-distribution plots do not agree with data. Even with the elimination of candidates under 2 MeV from the only simulation there is an over simulation of candidates in the forward direction for both slices of q3. The candidates are removed from all bins instead of targeting the forward direction where the over simulation is seen. Though simulation does move in the direction of data.

The next set of plots in Figure 7.7 are made while the algorithm randomly eliminates 12 percent of candidates under 10 MeV. To do this all candidates under 10 MeV were assigned a random number between 1-100 such that probability of each number from 1-100 being assigned is equal. Anything with a number from 1-12 was ignored by the algorithm. Now the plots are made in comparison to the data with the 2 MeV candidate threshold cut being applied. Once again the hadronic energy plots do not change.

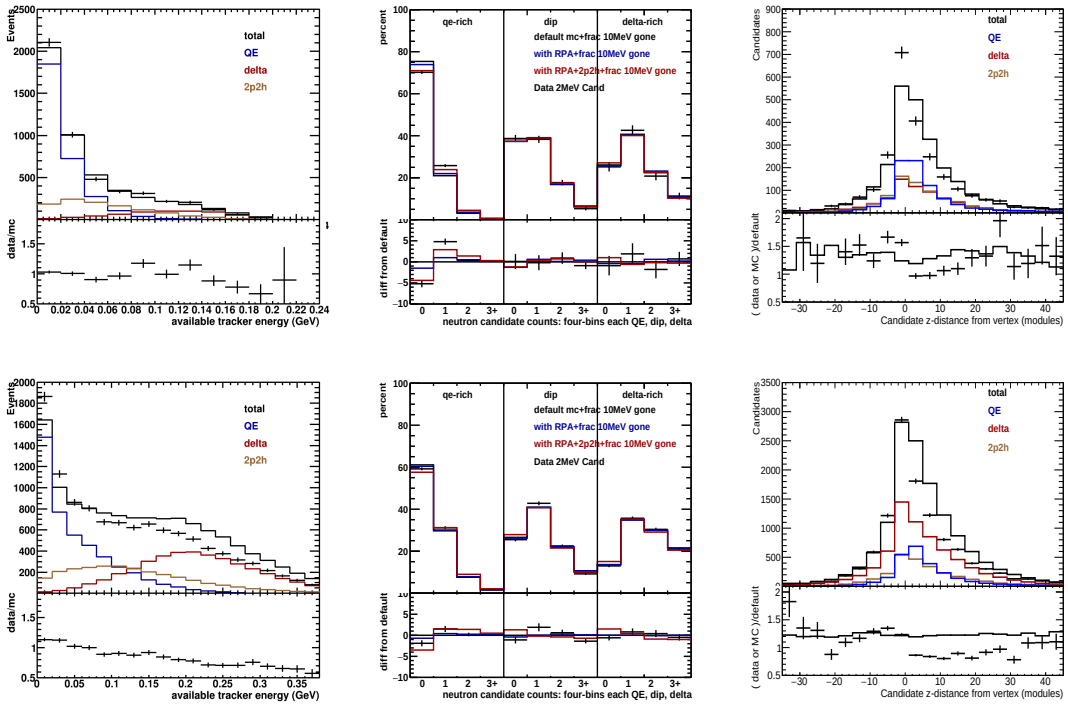

Figure 7.7: Simulation with RPA, 2p2h, 2 MeV candidate cut, and 12 percent of candidates under 10 MeV eliminated and data with a 2 MeV candidate cut. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the default simulation with a 2 MeV candidate energy threshold plus the 12 percent elimination while the ratio for the left plots is to the current simulation.

The neutron counting plots displayed here show how many candidates under 10 MeV need to be eliminated from the simulation so that it will describe data given the data also has the 2 MeV candidate cut. Removing 12 percent of the candidates causes the simulation to match data in most bins for both the low q3 and mid q3. With systematic error bars these would all describe data. The z-distribution plots still have an over simulation of candidates in the forward direction. Since the candidates under 10 MeV are eliminated randomly, once again they are eliminated from each bin causing the decrease to appear everywhere instead of targeting the forward direction.

The next set of plots are in Figure 7.8 and include the same configuration as the last set except now 24 percent of candidates under 10 MeV are eliminated. Since the previous set described data so well for the neutron counting this set now has simulation moving away from the data. The elimination of even more candidates causes the percentage of events with candidates to now be under simulated. The

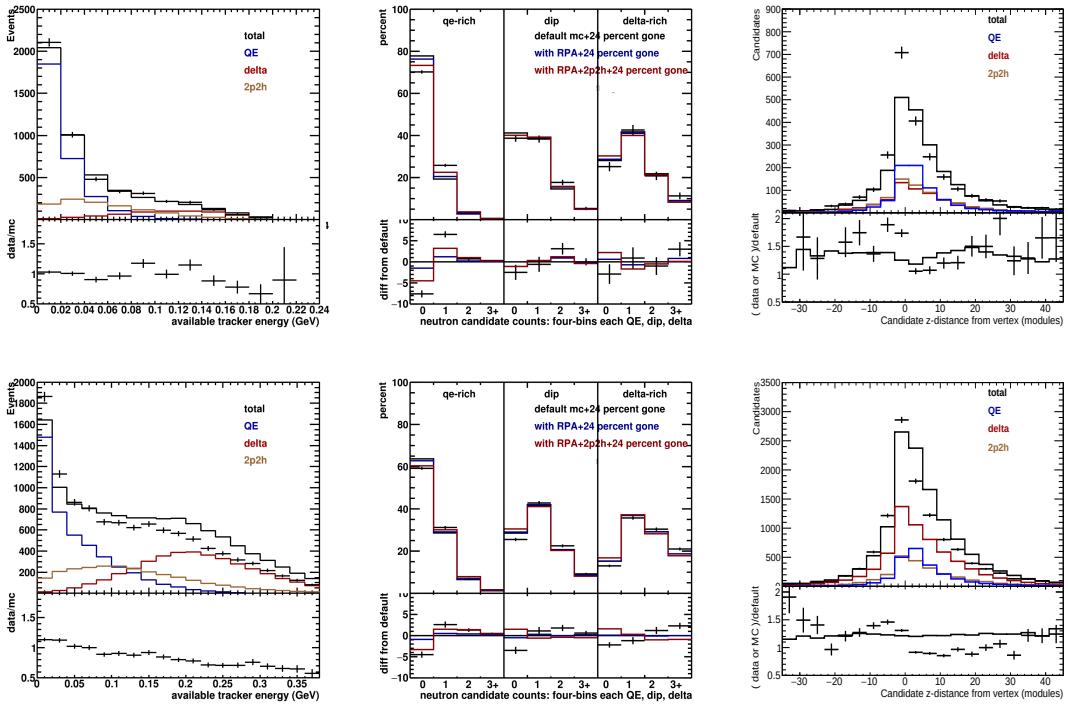

Figure 7.8: Simulation with RPA, 2p2h, 2 MeV candidate cut, and 24 percent of candidates under 10 MeV eliminated and data with a 2 MeV candidate cut. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the default simulation with a 2 MeV candidate energy threshold while the ratio for the left plots is to the current simulation.

z-distribution plots once again lose candidates from all bins instead of targeting the forward direction and thus the over simulation still exists here.

The final set of plots are displayed in Figure 7.9. This set shows how the z-distribution and neutron counting plots would change if the 24% of candidates less than or equal to 10 MeV eliminated were only those caused by GENIE neutrons and compares this MC to the default data. The neutron counting plots do not agree between simulation and data as shown previously. The z-distribution plots still show a decreases in all of the bins but now the forward direction is targeted more often than in the previous sets of plots. This is because the forward direction of this histogram contains approximately two-thirds of the candidates caused by GENIE neutrons.

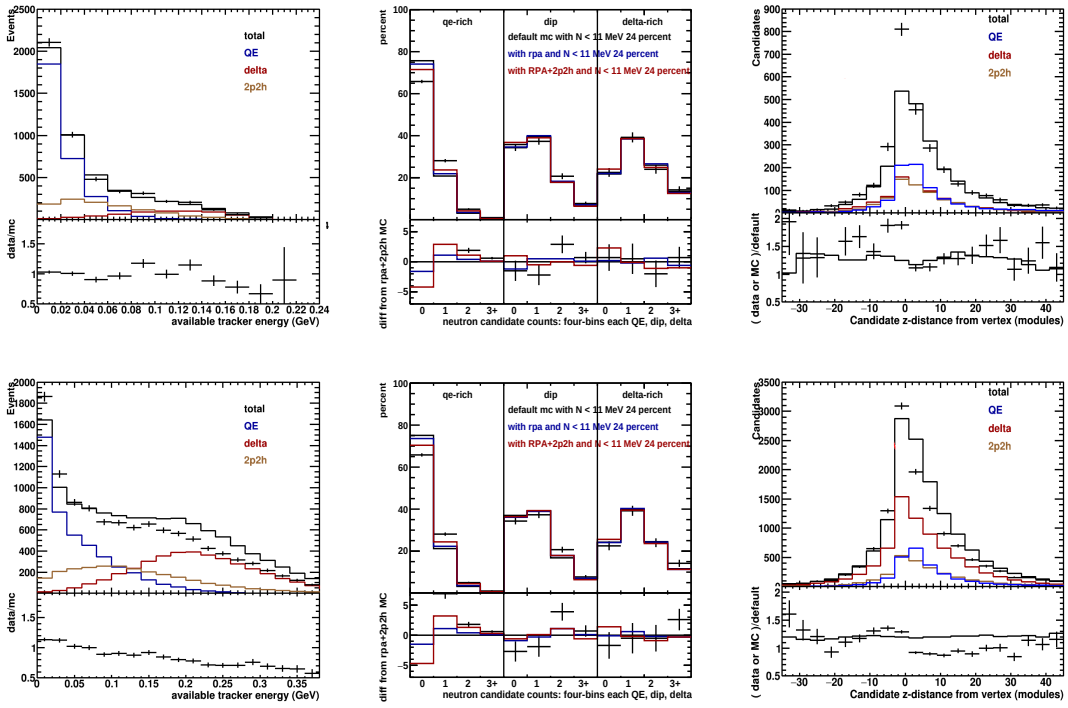

Figure 7.9: Simulation with RPA, 2p2h, 2 MeV candidate cut, and 24 percent of GENIE neutron caused candidates under 10 MeV eliminated and default data. Plots for the low q3 (top) and mid q3 (bottom). Neutron counting plots (middle) include data, the default simulation, default plus RPA, and current. The z-distance module plots (right) show the data and current simulation (black) broken down into QE events (blue), 2p2h (brown) and delta events(red). The hadronic energy plots for data and simulation separated by interaction type (left). Ratio for the right two sets of plots is to the default simulation with a 2 MeV candidate energy threshold while the ratio for the left plots is to the current simulation.

Chapter 8

Results and Conclusion

This thesis is an analysis of the MINERvA collaboration's anti-neutrino data and simulation to determine whether or not certain nuclear effects cause better agreement between the two. Given neutrons make up a significant portion of the products in the anti-neutrino interactions in this analysis a neutron counting algorithm was created to count the total number of neutrons there were present in either simulated or data events. Various cross section and model tweaks were investigated and then distributions created with the neutron counting algorithm were compared for both data and the tweaked simulations.

The result is the current and modified simulations do not simulate the generation and propagation of neutrons through the MINERvA detector well. The z-distribution plots show an excess of simulated neutron candidates in the forward direction and is not helped by any model or systematic effect tested.

The addition of the RPA suppression decreased the number of neutron-less events in the simulation for the QE-rich hadronic energy region neutron counts, pulling simulation towards data in this distribution, indicating a need for this type of effect.

However, there is a definite over simulation of neutrons in comparison to the data noticed from both the neutron counting and z-distribution plots, so the addition of the 2p2h process only increased the difference in the number of neutrons there are per event. The studies of eliminating random candidates did not affect the QE-rich region as much as the other regions and it retained some sensitivity to the 2p2h component. The addition of the 2p2h did improve agreement for the proton multiplicity comparisons. While not as prominent as neutrons, protons do tend to be products of the 2p2h and delta type anti-neutrino interactions analyzed here. The proton counting for this analysis tells the same story as in the neutrino case showing the inclusion of the 2p2h events helps agreement between data and simulation in both proton counting and in hadronic energy plots. It is likely the best description of the data includes a simulation with a combination of the RPA, 2p2h, and an additional systematic.

Only when the hadronic energy was increased by ten percent did the low q3 neutron candidate counts agree with the data and had an effect as large as the 2p2h on the mid q3 even if it did not cause agreement between the data and simulation. The hadronic energy agreement appeared better for the low q3 and mid q3 given the increase in hadronic energy caused events from lower slices of three-momentum

transfer to migrate to higher slices decreasing the excess of simulated events present here.

The over simulation of neutrons could also be fundamental to GENIE's FSI model, which is known to produce an excess of neutrons compared to NuWro and NEUT. The FSI model is applied to all types of reactions, and this feature may not be explicitly covered by the intra-nuclear re-scattering uncertainties.

Bibliography

[1] J. Nieves J. E. Amaro and M. Valverde. "Inclusive quasielastic charged-current neutrino-nucleus reactions". In: *Physical Review Letters* 70.5 (2004). DOI: https://journals.aps.org/prc/abstract/10.1103/PhysRevC.70.055503.

[2] J. Nieves I. Ruiz Simo and M. J. Vicente Vacas. "Inclusive charged-current neutrino-nucleus reactions". In: *Physical Review Letters* 83.4 (2011). DOI: https://journals.aps.org/prc/abstract/10.1103/PhysRevC.83.045501.

[3] R. Gran J. Nieves F. Sanchez and M. J. Vicente Vacas. "Neutrino-nucleus quasi-elastic and 2p2h interactions up to 10 GeV". In: *Physical Review Letters* 88.11 (2013). DOI: https://journals.aps.org/prd/abstract/10.1103/PhysRevD.88.113007.

[4] et al L. Aliaga. "Design, Calibration, and Performance of the MINERvA Detector". In: *Science Direct* 743.11 (2014), pp. 130–159. DOI: http://dx.doi.org/10.1016/j.nima.2013.12.053.

[5] et al. Aaron S. Meyer. "Deuterium target data for precision neutrino-nucleus cross sections". In: *arXiv* (2016). DOI: https://arxiv.org/pdf/1603.03048.pdf.

[6] Emily Draeger. "Proton Therapy Seminar". In: (2016).

[7] et al P. A. Rodrigues. "Identification of Nuclear Effects in Neutrino-Carbon Interactions at Low Three-Momentum Transfer". In: *Physical Review Letters* 116.7 (2016), pp. 1–6. DOI: 10.1103/PhysRevLett.116.071802.

[8] Chris Marshall Joshua Hignight Jake Calcutt Kendall Mahn. "Neutron Multiplicity". In: (2017).

[9] John Demgen. *Nuclear Effects in Quasi-Elastic and Delta Resonance Production at Low Momentum Transfer*. URL: http://minerva-docdb.fnal.gov:8080/cgi-bin/RetrieveFile?docid=11370&filename=John%20Demgen%20Thesis.pdf&version=1.

[10] Jeffery Kleykamp. *Hadron Reweight Tool*. URL: http://minerva-docdb.fnal.gov:8080/cgi-bin/RetrieveFile?docid=10689&filename=Kleykamp_2015-2-11.pdf&version=1.

[11] Tony Mann. *Isospin relations among CC(Pi) channels*. URL: http://minerva-docdb.fnal.gov:8080/cgi-bin/ShowDocument.

[12] Ethan Miletenberger. *Nuclear Effects in Neutrino Interactions at Low Momentum Transfer*. URL: http://minerva-docdb.fnal.gov:8080/cgi-bin/RetrieveFile?docid=11208&filename=EthanMiltenbergersThesis.pdf&version=1.

[13] Lu Ren. *Muon Fuzz Study*. URL: http://minerva-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=10340.

[14] Nathaniel Tagg. *Arachne: A Web-Based Event Viewer for MINERvA*. URL: http://minerva-docdb.fnal.gov:8080/cgi-bin/RetrieveFile?docid=3826&filename=Arachne%20-%20first%20presentation.pdf&version=3.

Appendix A

Hadron Re-Weights

One of the largest contributors to the neutron counting discrepancies seen in data and simulation comparisons was expected to be the uncertainty with the cross sections and mean free paths of the hadrons produced in the anti-neutrino interactions, specifically the neutrons. To help understand this uncertainty a systematic developed by MINERvA collaborator Jeffery Kleykamp, nicknamed the Hadron Re-weight Tool, was used to mimic changes in these cross sections and then both neutron and proton counting was completed [10].

A.1 Outline of the Systematic

The main idea of this systematic is to mimic a change in the cross section for various particles that could be present in the final state of neutrino and anti-neutrino interactions. In general a cross section is the effective area for which an interaction would occur in. In high energy physics the term cross section is also used to refer to the probability that a particle will scatter with another and react in a certain way. The are two types of cross sections that can be discussed as well. Cross sections that are a function of some variable such as scattering angle or energy, as seen with neutrons and protons, are known as "differential cross sections". One can integrate the cross section over all such variables to get the "total cross section". The difference between these two is important and should be noted.

The particles varied in the Hadron Re-weight Tool (which are protons, charged pions, and neutrons) have differential cross sections that are functions of energy. This is easily seen in Chapter 4 from the discussion of the dependence of a neutron's visibility based upon its energy. Thus the total cross section is proportional to the integral of the differential cross section with respect to the energy. The equation used in this systematic can be seen in A.1. The cross section for each of the particles is different for the material it will be interacting with, due to the number of nucleons present in the material, and thus this is calculated for neutrons,protons, and pions based on whether they are interacting with the carbon, iron, or lead nuclei within the MINERvA detector. Pions are included in this analysis due to the delta type

events having their delta particles decay into nucleons and a pion.

$$\sigma = \frac{1}{E_f - E_i} \int_{E_i}^{E_f} \sigma(E) dE \tag{A.1}$$

The average distance a particle will travel between separate collisions with another specified particle is known as its mean free path. This quantity is inversely related to the total cross section and the number of the type of particle being scattered off of. In this systematic the weight of events with a specified particle type is changed based on how far the particles of the specified type in the given event travel between interactions. Weighting events up with a given particle that travels far distances between interactions, or has a long path length, essentially is weighting down the cross section for that particle. This is easily noticed because we now have more events with this particle type moving far inside the detector. This increases the average mean free path and essentially decreases the cross section.

The weights calculated for this systematic are based upon the cross section, density, path length, and energy of the particle type chosen within the interaction. The user must decide whether or not to weight up or down the cross section of the chosen particle type. As an example, if the user wants the cross section to be weighted down the tool will find particles that travel a far distance, since these are examples of low cross sections and large mean free paths, and weight them up. If the particle travels a short distance, an example of a high cross section and small mean free path, it is weighted down.

The equations to calculate the weights used in the Hadron Re-weight tool are seen in equations A.2 to A.3. The first two equations are used for in-elastic interactions while the last is used for elastic, charge exchange, and non-existent (did not interact) type interactions. Once a weight is calculated for each interacting particle these are summed and a re-normalization factor is created by dividing the number of events by the summed weight. Each individual weight for every event, created by multiplying the weights together for each interacting particle, is multiplied by this re-normalization factor to insure the preservation of probability.

$$W_{int,up} = \frac{1 - e^{-x\rho\sigma(1+\delta)}}{1 - e^{-x\rho\sigma}} \tag{A.2}$$

$$W_{int,down} = \frac{1 - e^{-x\rho\sigma(1-\delta)}}{1 - e^{-x\rho\sigma}} \tag{A.3}$$

$$W_{non-int} = e^{-x\rho\sigma\delta} \tag{A.4}$$

The value for delta is dependent upon what type of material the particle interacted in. The MINERvA detector is made up of carbon, iron, and lead and the delta values for them are 0.1, 0.15, and 0.2 respectively. These values are approximately the uncertainty of the cross sections within the material and approximately what percentage the cross sections will be adjusted by. This is true for all of the particles except for neutrons while interacting within carbon nuclei. Neutrons are less understood then both pions and protons and so a separate value of delta is determined for

them interacting within carbon as it is the most likely place for them to do as such. It is known that the neutron's cross section is highly dependent upon its kinetic energy so the delta value here is calculated as a function of the kinetic energy of the neutron taking part in the interaction. These delta values tend to be larger then the 0.1 for the proton and pion interacting with carbon but never larger then 0.25.

To test the Hadron Re-weight tool, Jeffery used another tool called a particle cannon was used. This tool allows the user to simulate any given number of a specified particle interacting within the MINERvA detector. The particle cannon was used to simulate 10,000 protons with energies between 0.5 GeV and 5 GeV and 10,000 pions with 0.75 GeV of energy. The path lengths of each of these particles were saved given the default model and then were saved again with the mean free paths being weighted both up and down. These were then compared to make sure the interaction path lengths, as this is what is changed to mimic the mean free path and cross section changes, of the weighted events were not more then approximately 10 percent different from those in the default simulation.

Another test done was just to ensure that the weights for each event made sense. Given the particle cannon was used to simulate protons and pions, each event should have had one of these two within. That means each event should have had a weight that was not equal to one when the Hadron Re-weight Tool was turned on. To check this a histogram of the weights for every simulated particle cannon event was created for both protons and pions. When zoomed in near one it was apparent there were no events with a weight of one. Since the interaction path lengths were adjusted reasonably given the parameters assigned to the weight equations and the weights themselves were as expected the tool is believed to work.

The exact checks used by Jeffery could not be done here due to the variables saved in the anti-neutrino simulation data used for this study. However, it was possible to plot the weights for every event and the weight of the particle by distance traveled. The plots for the low q3 are visible in Figure A.1.

The weights for the neutron cross section changes after re-normalization for the low q3 are seen on the left in this figure and only contain the events that had a neutron. Those that did not would get an event weight of 1 because no change would be made to a particles cross section in the event. The event weight plots have a mean of 0.999 (top) and 0.996 (bottom) which is near the expected average of 1. This average should be 1 to ensure the overall event rate did not change. Because these events all contain neutrons none of the weights in these plots are 1. It was check that they all have a small deviation if they fall within the bin containing 1.

The plots of the right show the weight given to each neutron in every event in the low q3 versus the distance the neutron traveled before it interacted or left the detector and did not interact. Remember this is not the same as the event weight because the event weight is created by multiplying the weight given to each particle with a cross section change and then multiplying this by the re-normalization factor. However this plot helps ease the mind when determining if the Hadron Re-weight code works because the weight should decrease and distance increases when the cross section is being weighted up and the opposite when the cross section is being weighted down. This is exactly what is seen.

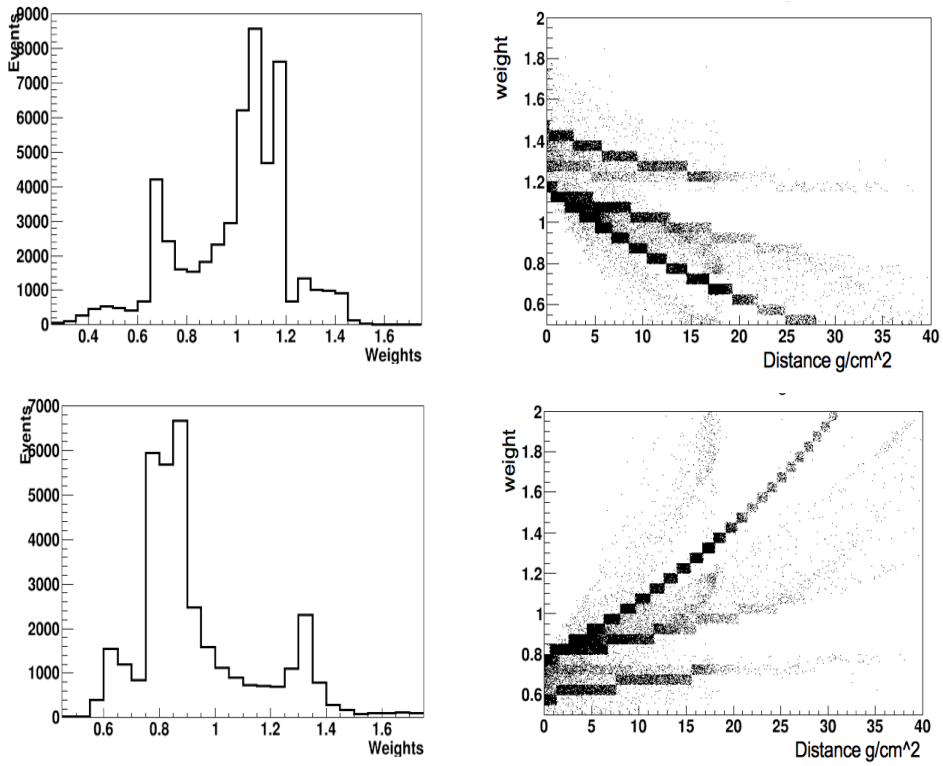

Figure A.1: Top: Plots for weighting the cross section up. Bottom: Plots for weighting the cross section down. Left: Weight for events with only neutrons in them. Right: Weight for each neutron in low q3 by the distance it traveled. Distance is measured in units of g/cm^2.

The different black lines represent the different interaction types that the neutrons are going through with the two darkest lines, bottom 2 for the top plot and opposite for the bottom plot, being completely non-interacted and elastic type interacted neutrons. These two interactions both fall within the non-interacted category and weights are calculated using A.4. It is important to note that this gives many events with weights less then one when the cross section is weighted up and just as many events with a weight greater then one when the cross section is weighted down. This causes the change in the neutron candidate counts to be counter-intuitive.

A.2 Neutron Counting

Neutron counting was completed using the Hadron Re-weight Tool by increasing and decreasing the cross section of protons, neutrons, and pions by 1 sigma. Each of these effects were done separately and impose an additional weight to the simulation on top of the RPA and 2p2h effects which are included as well. The plots created for this part of the analysis are once again the same as those seen in the previous neutron and proton counting sections.

The cross section changes that should affect the neutron counts the most are obviously those associated with the neutrons themselves. These changes are associated

with a one sigma adjustment which corresponds to approximately the value for the delta in Equation A.2 to Equation A.4 and depends on the material the particles interacted in. The plots for mimicking an increased cross section are shown in Figure A.2. Due to the interaction types that are most present most of the weights for the particles will be less then one so this should actually cause a decrease in the number of neutron candidates counted and is exactly what is seen. This is because most of the weights being less then one causes the re-normalization factor to be greater then one. Once the event weights get multiplied by this the events that don't interact, which show up in the zero candidate bin and are a majority, end up increasing in the overall percentage. While the changes are similar between the low q3 and the mid q3 they are not between hadronic energy regions.

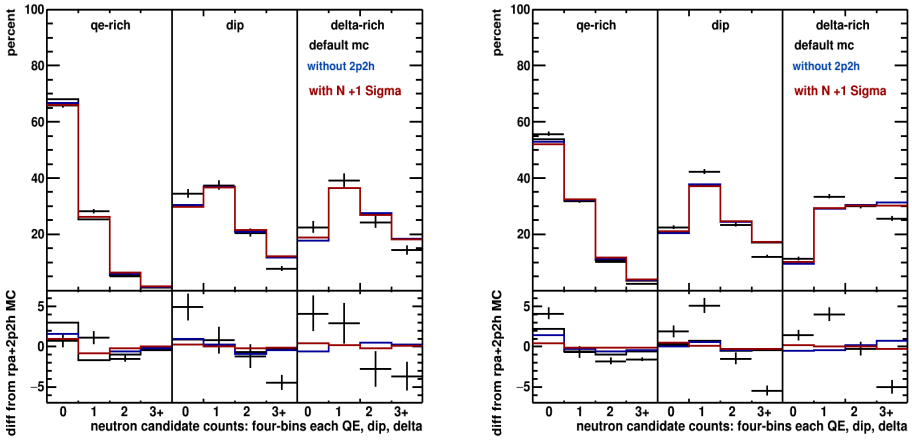

Figure A.2: Percentage of neutron candidates per event while weighting the neutron cross section up for the low q3 (left) and mid q3 (right) for the QE, dip, and delta regions.

The QE hadronic energy region is almost completely all QE type events which always have a neutron in the final state. These neutrons are visible at most half of the time, mainly due to the quasi-elastic neutron re-interactions that are counted as non-interacted. If the events with visible and invisible neutrons interacting in them are weighted down, but differently because the path length for the different interactions will cause differnt weights, the effect here should be very noticeable and is. There is an increase in the percentage of events with no candidates in them with a similar decrease in events with one candidate. There are very few events with more then one candidate here because of the one neutron final state probability in the QE region so the corresponding bins are not affected by a large change.

The dip and delta regions have slight increases in the percentage of events with zero candidates in them. The magnitude of the change is smaller in these two sections in comparison to the QE region. The dip region has a smaller change likely due to the two neutron final state. Here it is probable the events with two neutrons in the final state would have one of them visible and possibly two. Since nearly all visible neutrons interact in the same way this causes almost all of the neutrons to be weighted down and in similar ways. After re-normalization this causes only a slight

increase in the percentage of events with zero candidates. The delta region does not have two neutron final states but it does have energetic neutrons in the final state which are the most likely to be visible. If this is true for a majority of the events, and they are all quasi-elastic interactions calculated using the same non-interacting weight, then the same argument can be made here.

The plots for mimicking a decreased cross section show opposite effects as those increasing the cross section and are visible in Figure A.3. Here events with small interaction lengths and most that are within the non-interacted category are weighted up. Since most of the neutrons fall within the non-interacted category this causes a noticeable increase, with a larger magnitude then the opposite study, in the percentage of events with one or more neutron candidates in them. This larger change can possibly be attributed to the fact almost all of the events with neutrons in them are either truly non-interacting or have a quasi-elastic re-interaction categorized as non-interacting. If all of these are weighted up then the re-normalization factor is less then one here and causes this asymmetric change. Neither the neutron cross section weighted up or down helps simulation agree with data.

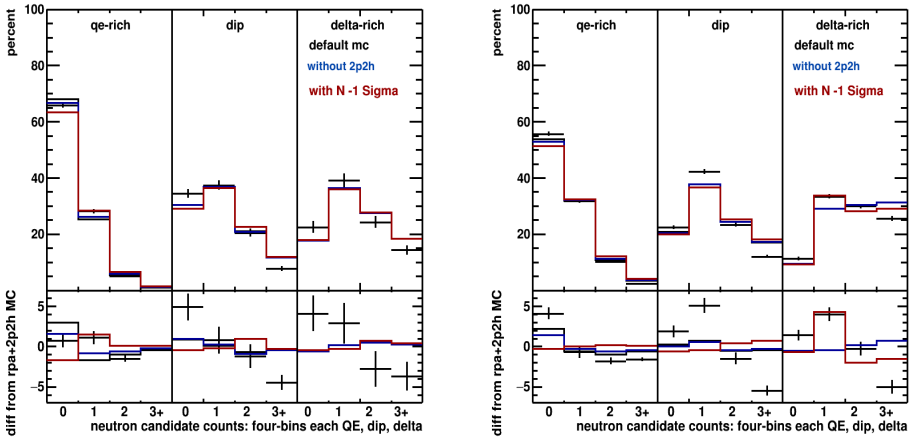

Figure A.3: Percentage of neutron candidates per event while weighting the neutron cross section down for the low q3 (left) and mid q3 (right) for the QE, dip, and delta regions.

The other weights added to the simulation from the Hadron Re-weight Tool included the increase and decrease by one sigma to the cross sections of protons and charged pions. Each of these changes were made individually and then neutron counting was done for every event once again. These modifications caused very little change in the total number of neutron counts there were.

Protons are a very small part of the background to the neutron counting for both the low q3 and the mid q3. Referencing Table 3.3 it can be seen that protons create about 0.28 percent of the neutron candidates in the low q3 and 1.54 percent in the mid q3. Given these percentages are so small any adjustment to the proton cross section, which maxes out at a 20 percent overall change to the cross section given the deltas provided, should not have a noticeable affect on the percentage of events with a given number of neutron candidates. Put another way is a 20 percent change

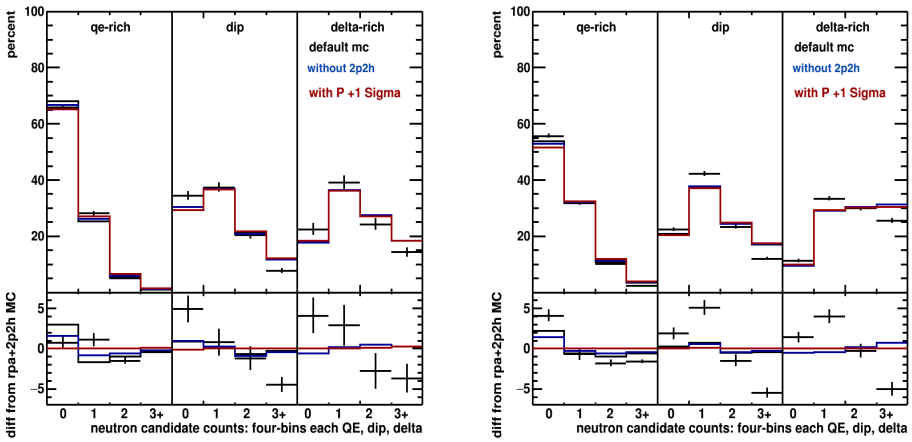

Figure A.4: Percentage of neutron candidates per event while weighting the proton cross section up for the low q3 (left) and mid q3 (right) for the QE, dip, and delta regions.

in the 0.28 or 1.54 percent of events with the proton caused neutron candidates in them is negligible, especially because these events do not populate just one bin in the neutron counting plots. The plots for the adjustment in the proton cross sections for the low q3 and mid q3 are displayed in Figure A.4 and Figure A.5. It is apparent that the adjustments to the proton cross sections did not make any change to the neutron counts as expected.

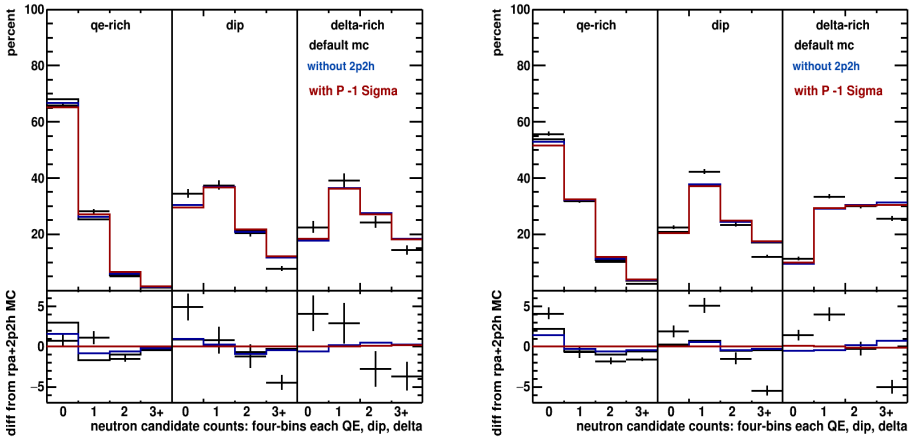

Figure A.5: Percentage of neutron candidates per event while weighting the proton cross section down for the low q3 (left) and mid q3 (right) for the QE, dip, and delta regions.

The pion background is much larger then that of the proton background as pions make 10.79 percent of the neutron candidates in the low q3 and 22.31 percent in the mid q3. However, even though this background is much higher the effect of weighting the cross section of the pions either up or down is once again negligible. As before the maximum value the cross section of the pion can change is by 20 percent in

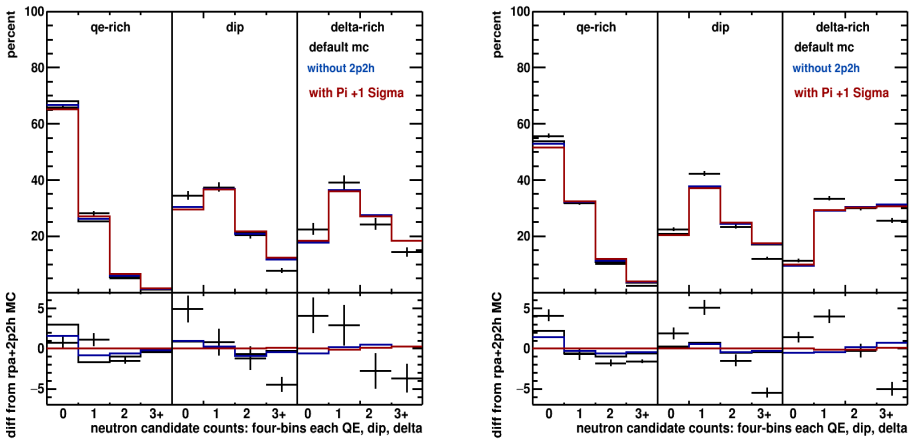

Figure A.6: Percentage of neutron candidates per event while weighting the pion cross section up for the low q3 (left) and mid q3 (right) for the QE, dip, and delta regions.

each event. This, however, is unlikely as the 20 percent change is associated with the pion interacting in lead which would put the interaction outside of the tracker part of the MINERvA detector. Most of the pions interact within the carbon nuclei and thus the change in the cross section is approximately 10 percent. A 10 percent change in 10.79 percent of the candidates in the low q3 and even 22.31 percent of the candidates in the mid q3 still gives a small change. When the events with these pion caused candidates are spread out over multiple bins in the dip and delta regions the overall percentage of events with a certain number of candidates show minimal change.

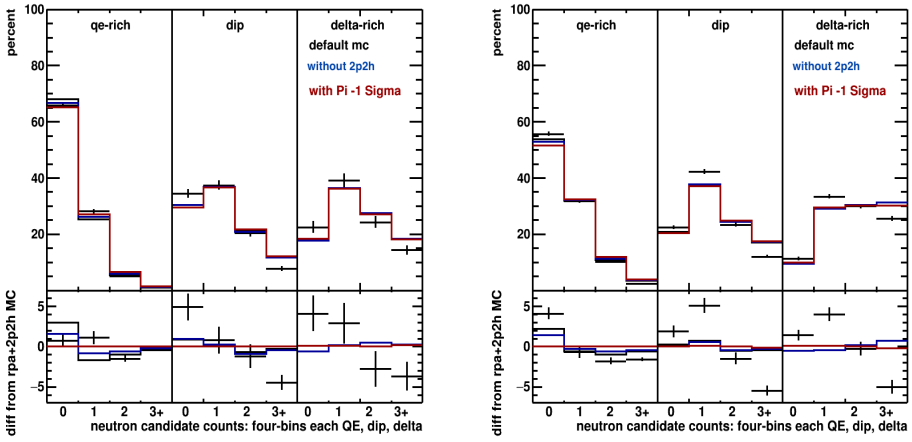

Figure A.7: Percentage of neutron candidates per event while weighting the pion cross section down for the low q3 (left) and mid q3 (right) for the QE, dip, and delta regions.

The plots showing the neutron candidate counts while changing the pion cross section up can be seen in Figure A.6. Here pions would have very short mean free

paths and would get an increase in their event weight. This causes a slight increase in neutron candidates which appears in the delta regions for the low q3 and the mid q3. This makes sense due to the delta interaction, which has the pion in the final state, populating these regions. The plots showing the decrease in the pion cross section appear in Figure A.7 and shows an equal in magnitude and opposite effect.

A.3 Proton Counting

Proton counting was completed using the Hadron Re-weight tool by increasing and decreasing the cross sections of protons, neutrons, and pions by 1 sigma. This study gives an idea of the contamination level of neutrons and pions within the proton candidates, on top of the neutron causing proton candidate study shown in Chapter 5. The tables shown below once again follow the same scheme as before. The logic laid out for most of these studies from the neutron counting follows through to the proton counting as well.

It would be assumed if the cross section of neutrons interacting was either increased or decreased this should cause no effect on the proton candidates in each event. However, there is a small change seen in the percentage of proton candidates per event, shown in Figures A.8 and A.9, which helps identify there must be a larger background of neutrons causing proton candidates then there are of protons causing neutron candidates since the change of the proton cross section did not affect the neutron candidates.

Protons are well understood and it is known they will interact very near the anti-neutrino interaction vertex so the candidates are localized and searched for here. However, neutrons are not understood very well but are expected to interact further from the anti-neutrino interaction vertex more often then not so are not searched for in this area. If protons always interact near the vertex and neutron candidates are counted outside this area the proton background in neutron candidates would be small, and is. If neutron candidates usually do not interact near the vertex, but sometimes do, and proton candidates are counted here then there should be a background of neutrons in the proton candidates. This was discussed previously and shown in Figure 6.2 and Figure 6.3.

The neutron background in the proton candidates allows for an increase in the percentage of events with proton candidates when the neutron cross section is decreased which is shown in Figure A.9. The reasoning here for the counter-intuitive change is the same as discussed in the neutron counting section of this chapter since it is neutrons that are causing the proton candidates in the events changing here. This effect is more noticeable in the mid q3 then in the low q3 and grows with increasing hadronic energy as well. This would make sense as there would be a larger population of more energetic neutrons in the mid q3 and the dip and delta hadronic energy regions then in the QE hadronic energy regions. The neutrons causing proton candidates would have to be energetic enough to not only interact and do so near the vertex but also produce a 20 plus MeV cluster within the MINERvA detector.

The number of protons per event decreases slightly for both the low q3 and mid q3 when the neutron cross section is increased. The plots for this is seen in

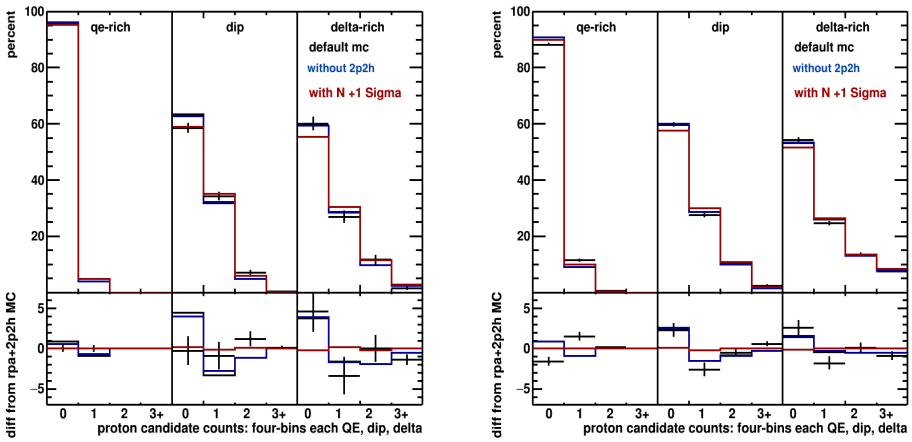

Figure A.8: Percentage of proton candidates per event while weighting the neutron cross section up for the low q3 (left) and mid q3 (right) for the QE, dip, and delta regions.

Figure A.8. The magnitude of the effect coming from the increase in the neutron cross section is smaller then that coming from the decrease of the neutron mean free path. This could be due to the imbalance between the number of neutrons either not interacting and interacting very far from the detector and doing so quasi-elastically, putting them into the non-interacting category for the tool, and those that interact near the vertex. It is important to remember this change only affects a very small number of proton candidates and thus the overall effect is now smaller then what was seen for the neutron counting case.

It's becoming clear that not everything follows what one would intuitively expect for the neutron or proton counting and the proton counting with the proton cross section changes follows this trend. It is important to remember that protons are well understood and it is expected that the simulation does in fact simulate the proton interactions well. All of the proton candidates counted within the MINERvA data and simulated events fall within 80mm of the anti-neutrino interaction vertex. This means they all have a similar interaction length and thus re-weighting these to mimic changes in the cross section does very little after re-normalization. The plots for the proton cross section weighted up are in Figure A.10 and Figure A.11.

The proton counting plots for the mimic of the increase in the proton cross section with various materials show exactly what is discussed in the previous paragraph. There are no changes noticeable to the percentage of events with a given number of proton candidates counted within. This remains true for the decrease in the proton cross section.

As with the previous hadron re-weighting for the proton counting case there is not a large change in the overall percentage of events with a given number of proton candidates when the pion cross section is adjusted. The increase of the pion cross section should make pions more visible while the decrease should make pions less visible, forcing the pion caused proton candidates visibility to follow the same pattern. The pion cross section change up plots are shown in Figure A.12 and the

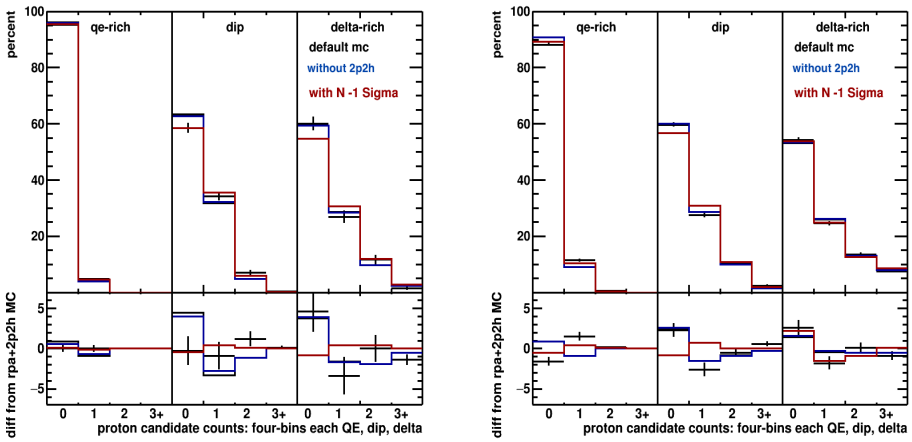

Figure A.9: Percentage of proton candidates per event while weighting the neutron cross section down for the low q3 (left) and mid q3 (right) for the QE, dip, and delta regions.

cross section down plots are in Figure A.13.

It is very rare for a pion to be visible in the QE region and, while less rare, is still not common in the dip region as well. This idea agrees with what is shown in the proton counting plots for either q3 slice. There is no noticeable change in the QE region for either the low q3 or the mid q3 and the changes in the dip region are less then half of a percent for either the pion cross section increase or decrease. The only noticeable and notable changes are in the delta regions.

The delta region is nicknamed as such due to the large population of delta type events that fall within it. These events are known to produce delta particles which quickly decay into a nucleon plus a pion. This would indicate if there was a change in the proton counting due to a pion contamination in the proton candidates it would appear in this hadronic energy region and it does. In the proton counting algorithm any clusters of energy caused by a particle that is tracked are ignored. For the case of pions most of these particles do fall into the category of tracked with the exception of a few that produce a cluster just off of the track. This means a significant portion of the pions produced in the delta region do not meet the requirements to be a proton candidate and while this region of hadronic energy is where any changes in proton counting would be seen the change should still be small. This type of change is exactly what is seen and follow the trends expected. The increase of the pion cross section causes a increase in events with 1 or more proton candidates in them. The opposite is true for the decrease of the pion mean free path or increase of the pion cross section. However this once again does not help agreement between data and simulation.

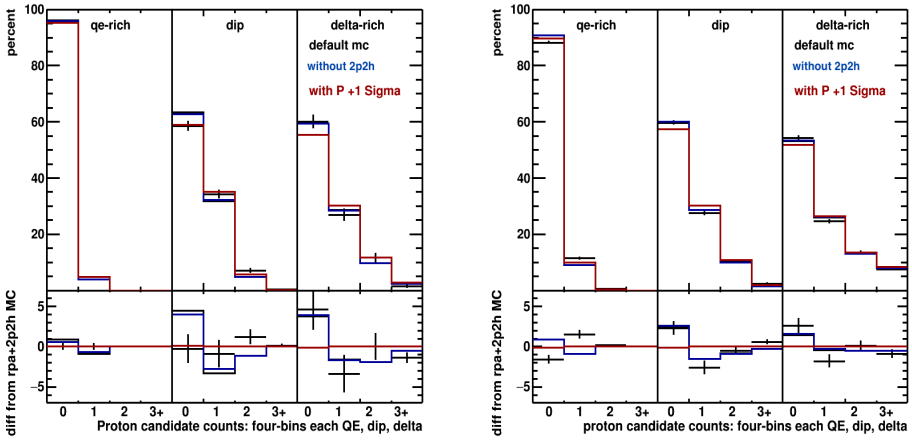

Figure A.10: Percentage of proton candidates per event while weighting the proton cross section up for the low q3 (left) and mid q3 (right) for the QE, dip, and delta regions.

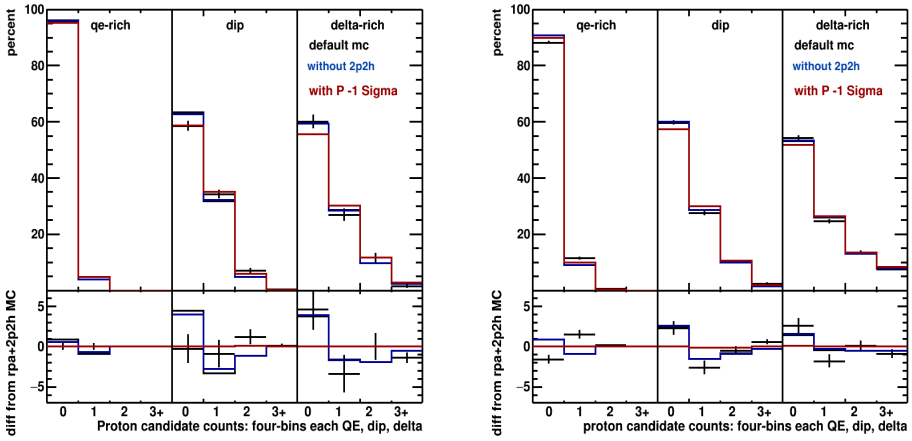

Figure A.11: Percentage of proton candidates per event while weighting the proton cross section down for the low q3 (left) and mid q3 (right) for the QE, dip, and delta regions.

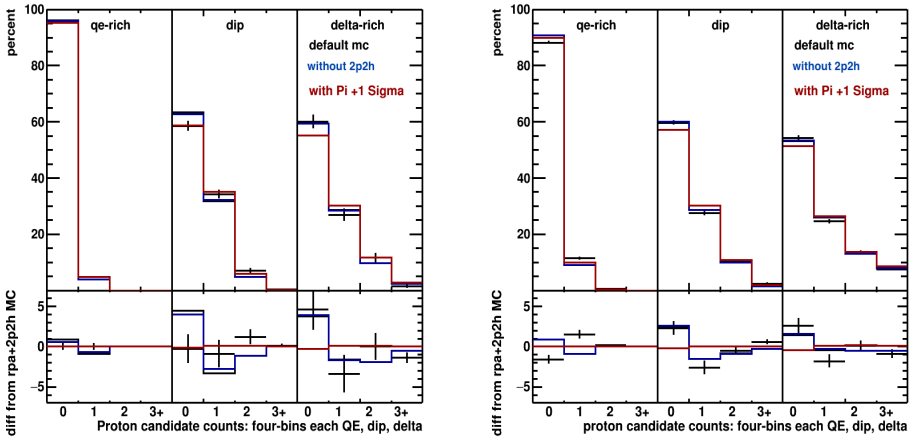

Figure A.12: Percentage of proton candidates per event while weighting the pion cross section up for the low q3 (left) and mid q3 (right) for the QE, dip, and delta regions.

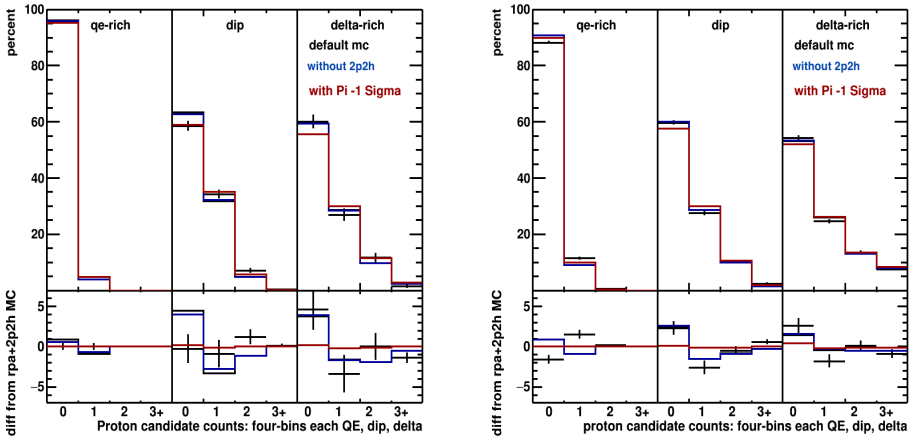

Figure A.13: Percentage of proton candidates per event while weighting the pion cross section down for the low q3 (left) and mid q3 (right) for the QE, dip, and delta regions.